Essays in quantitative finance

Akyildirim, Erdinc

Essays in Quantitative Finance

Dissertation
for the Faculty of Economics, Business Administration and Information Technology of the University of Zurich

to achieve the title of
Doctor of Philosophy
in Banking & Finance

presented by

Erdinç Akyıldırım

from Turkey

approved in October 2013 at the request of

Prof. Dr. E. Walter Farkas
Prof. Dr. H. Mete Soner

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorizes the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zurich, 23 October 2013.

The Chairman of the Doctoral Committee: Prof. Dr. Josef Zweimüller.

To *Sara, Nejdet, Berrin, Öznur.*

Acknowledgements

First and foremost, I would like to express my immense debt of gratitude together with heartfelt thanks to Prof. Dr. H. Mete Soner for his excellent guidance, infinite patience, and the considerable amount of advice that, all together, kept me on the right track to complete my Ph.D. studies. I always felt privileged to have the opportunity to work with him. I could not have asked for a better role model who not only exemplifies excellence in research but also in personality.

I am indebted to Prof. Dr. E. Walter Farkas for his invaluable recommendations and helpful comments on my work. I owe special thanks to Prof. Dr. Jean-Charles Rochet for attending my thesis defense and for his contributions to our joint paper. I would like to extend my special thanks to Prof. Dr. Vedat Akgiray, Prof. Dr. Alp Eden, Prof. Dr. Ali Rana Atılgan, Dr. Muzaffer Akat and Dr. Feyzullah Eğriboyun for their warm encouragement and constant support ever since the early stages of my Ph.D. studies. I am also very grateful to Dr. Emrah Şener and Prof. Dr. Loriano Mancini for the opportunity to work with them on very interesting topics and to learn from them. Many thanks are also due to Denise Künzli for her endless patience and kindness on administrative issues.

I will forever be thankful to İ. Ethem Güney, Dr. Selim Gökay, Albert Altarovici, and Dr. Rafaela Guberovic for always being next to me whenever I needed their friendship and help. I believe that our friendship has extended well beyond our shared time in Zurich. I owe very special thanks to Nilüfer Çalışkan, Ozan Akdoğan, Elise Gourier, and Robert Huitema for smiling faces whenever I knocked on their door. I would also like to express my thanks to my office mates Edgars Jakobsons, Ren Liu, Mario Sikic and Danijel Zivoi for creating a warm and joyful working environment at the Mathematics Department of ETH Zurich. Furthermore, I will always remember my friends Chris Bardgett, Mustafa Karaman, Felix Matthys, Cosimo-Andrea Munari, Alper Odabaşıoğlu, Tatjana-Xenia Puhan, Annina Saluz, Mirjana Vukelja.

I gratefully acknowledge financial support from the European Research Council under the grant 228053-FiRM and also from the ETH Foundation for making this thesis possible.

Above and beyond all, there are not enough words to express my deepest gratitude and love to my parents Sara and Nejdet Akyıldırım and my sister Berrin for their dedication, support, and patience in every possible way. I am also indebted to my wife, Öznur, for her continuous support, sacrifices, and love beyond any measure all of which allowed me to complete my Ph.D. journey.

Zurich, August 2013 Erdinç Akyıldırım

Contents

Chapter 1

Introduction and Summary of Research Results

1. Introduction

This dissertation deals mainly with the correct pricing and alternative hedging methods of financial instruments so called derivatives in both discrete and continuous-time under complete and incomplete financial market frameworks and also with the optimal dividend policy problem for firms. In particular, we first study the pricing of American and European options including the standard put and calls, barrier, lookback, and Asian type pay-offs under stochastic volatility by using a recombinant tree approach. Second, we analyse, from both empirical and theoretical perspectives, the pricing of dual currency credit default swaps in European markets, which in effect helps us to understand the currency dependence of these products. Thirdly, we investigate the problem of quantile hedging which is an alternative hedging strategy when an investor is unwilling or unable to put up the initial amount of capital required by a perfect or super-hedging strategy. Last but not least, we also examine the effects of interest rate fluctuations on the optimal dividend policies of firms when the drift and volatility of the cash flow process are constant.

2. Summary of Research Results

The thesis is based on the following four research articles:

(i) Approximating Stochastic Volatility by Recombinant Trees (with Yan Dolinsky and H. Mete Soner),

(ii) Dual Currency Credit Default Swap: Theoretical and Empirical Analysis (with Loriano Mancini and Emrah Şener),

(iii) Partial Hedging and Cash Requirement in Discrete Time,

(iv) Optimal Dividend Policy with Random Interest Rates (with İ. Ethem Güney, Jean-Charles Rochet, and H. Mete Soner).

Their content and contribution are summarized in the following subsections.

2.1 Approximating Stochastic Volatility by Recombinant Trees

In the literature, stochastic volatility models have been introduced to address the volatility smiles observed in option markets and the heavy tails and high peaks of the underlying asset distributions. The first research article of this dissertation develops a general approach to construct recombinant tree approximations for stochastic volatility models. This approximation as the Cox, Ross & Rubinstein (CRR) model easily constructs a discrete time financial market that itself is arbitrage free and as such allows for simple analysis of related complex instruments. Our main tool is to apply correlated random walks in order to approximate diffusion processes. A correlated random walk is a generalized random walk in the sense that the increments are not identically and independently distributed, but they only satisfy some Markov type of conditions. The resulting approximation is a four tuple Markov process. The first two components are related to the stock and volatility processes and take values in a two dimensional Binomial tree. The other two components of the Markov process are the increments of random walks with simple values in $\{-1, +1\}$. These processes naturally lie on a grid, and their Markov structure allows for an efficient computation of option prices. Our extensive numerical experimentation with the resulting pricing equations confirms the efficiency of the method. In the literature, tree based methods have also been considered. However, our approach differs from these earlier studies in two fundamental points. Firstly, our approximation is recombinant by construction while in the previous studies recombination is achieved through truncation. Also, our tree is arbitrage free and we provide a proof of convergence.

2.2 Dual Currency Credit Default Swap: Theoretical and Empirical Analysis

As markets remain uncertain over the debt dynamic sustainability of eurozone peripheral countries, sovereign CDSs are being widely adopted as an investment and risk management tool. Divergent investor demand for these contracts in USD vs EUR format has led to the evolution of an active quanto market in sovereign CDS. The second research article of this dissertation is an empirical study of the properties of violations of the law of one price in the European CDS market around period of market distress. We investigate the extent to which limits to arbitrage are state-dependent and how they have been affected during crises.

Our investigation is three-fold. First, we calculate a proxy of deviation from the LOP called *Quanto CDS*. This is based on the spreads between two CDS prices issued by the same sovereign but denominated in different currencies. In doing so, we take the correlation between default risk and exchange rate into consideration and form a variable that should be zero if the LOP holds. Then, considering the whole sample data (August 2010 to June 2013) which includes the Europe specific turbulence periods, we investigate whether our proxy steered and persistently fluctuated away from zero. Second, we conduct a factor analysis which is useful in reducing the dimension by concentrating on a few important factors that represent the main sources of variation in the dual currency CDS market. We find that the first two factors statically significant during the entire period of the time. Thirdly, in light of these results, we suggest and test a model in pricing dual currency CDSs in European markets, which in effect helps us to understand currency dependence. This is an important question since many financial institutions are tempted to employ USD as the *de facto* currency to price EUR denominated products of the same issuer. The motivation is that USD spread is a perfect substitute to discount risks denominated in other currencies given that risk premia (of the same issuer) across two currencies is identical, which obviously is not satisfied in segmented markets.

2.3 Partial Hedging and Cash Requirement in Discrete Time

In their seminal paper, Föllmer and Leukert [1] describe quantile hedging as the optimal hedge when the initial capital is less than the minimal super-hedging or perfect hedging cost. In particular, they determine the minimal amount of initial capital an investor can save by accepting a certain shortfall probability. Equivalently, they find the maximal probability of a successful hedge the investor can achieve if she is unwilling to put up the initial amount of capital required by a super-hedging (or a replication) strategy.

In a recent paper, Bouchard et al. [2] provide a different approach to quantile hedging. They consider the more general problem of finding the minimal initial data of a controlled process which guarantees reaching a controlled target with a given probability of success. As a special case, they focus on the quantile hedging problem and reproduce the explicit solution of [1] in continuous time. In the third research article of this dissertation, one of our goals is to further understand and develop their techniques in discrete time. We believe that our discrete time model has the advantage of streamlining the main ideas in [2] and bringing numerical difficulties to the surface. Our discrete time model can also recover the solutions for utility indifference pricing, good deal pricing, and expected shortfall, but our main contribution to the literature is in the context of quantile hedging.

There are a number of papers which study the quantile hedging problem in various frameworks. However, they require vastly different constructions under different market conditions. The main advantage of our discrete time model is that we can handle different market structures such as exotic options and markets with portfolio constraints by only slightly modifying our original method. The detailed numerical analysis of the quantile hedging problem in these varying frameworks verifies the efficiency of our dynamic programming equations and derived algorithms.

2.4 Optimal Dividend Policy with Random Interest Rates

There is a sizable literature on the optimal dividend policy problem for a company that is not allowed to issue new securities or obtain a new loan from a bank. In that case, the optimal dividend policy is simple and natural: distribute dividends whenever the level of cash reserves exceeds a certain threshold that depends on the characteristics (drift, volatility) of the cash flow process and the interest rate demanded by shareholders. An interesting extension of this problem is to investigate how the optimal dividend policy is modified when the profitability of the firm changes over time, due in particular to business cycle fluctuations. For example, Cadenillas and Sotomayor [4] solve for the optimal dividend policy when the drift and the volatility of the cash flow process are governed by a Markov chain representing macroeconomic fluctuations. Bolton, Chen &

Wang [3] study more generally the impact of changing macroeconomic conditions on both the financial and investment policies of the firms. However, Gertler and Hubbard [5] also show that macroeconomic conditions directly influence payments to shareholders, even independently of each firm's specific earnings performance. Two natural channels for this influence are the fluctuations in interest rates demanded by investors, and the conditions of the credit market.

The fourth research article of this dissertation examines how these macroeconomic fluctuations influence the dividend policies of firms, even in the absence of fluctuations in their earning processes. In other words, we study the polar case to the one considered in the literature: the drift and volatility of the cash flow process are constant, but the interest rate demanded by investors follows a Markov chain.

The fundamental economic result established by our paper is that the firm will distribute dividends more often when interest rates are high than when they are low. This result comes from the fact that the opportunity cost of cash reserves is higher when the interest rates demanded by investors are high. However, it does not fit well with the empirical evidence, given that firms actually tend to distribute less dividends during recessions (when interest rates are high) than during booms (when interest rates are low) even when the changes in firms' individual profitability are corrected for (Gertler and Hubbard [5]). This suggests that other macroeconomic factors, such as the size of frictions on financial markets, must play a role.This is why we also introduce the possibility for the firm to make new equity issuances. When the cost of these new issues (a proxy for the size of financial frictions) is substantially higher during recessions than during booms, the ranking of dividend thresholds is reversed, and firms now distribute more dividends during booms than during recessions. We also provide numerical evidence for the above conclusions.

Bibliography

[1] Föllmer, H. and Leukert, P. (1999). Quantile hedging. *Finance and Stochastics*, 3, 251–273.

[2] Bouchard, B., Elie, R. and Touzi, N. (2009). Stochastic Target Problems with Controlled Loss. *SIAM Journal on Control and Optimization*, 48(5), 3123–3150.

[3] Bolton, P., Chen, H. and Wang, N. (2011). Market Timing, Investment and Risk Management. *Working paper, Columbia University.*

[4] Cadenillas, A. and Sotomayor, L.R. (2008). Classical Singular and Impulse Stochastic Control for the Optimal Dividend Policy when there is Regime Switching. *Working paper, GSU and University of Alberta.*

[5] Gertler, M., and Hubbard, R.G. (1993). Corporate Financial Policy, Taxation and Macroeconomic Risk. *The RAND Journal of Economics*, 24(2), 286 – 303.

Chapter 2

Approximating Stochastic Volatility by Recombinant Trees

Erdinç Akyıldırım, Yan Dolinsky, and H. Mete Soner

This paper was accepted for publication in the **Annals of Applied Probability** (forthcoming).

2.1 Introduction

Contrary to many mathematical models, the discrete counterpart of the celebrated Black and Scholes [4] model came after its continuous version and it is generally accepted that this simple Binomial approximation by Cox, Ross, and Rubinstein [8] has been instrumental in the better understanding and the applicability of the model. Rubinstein [28] states that "the Black and Scholes model is widely viewed as one of the most successful in the social sciences and perhaps, including its binomial extension, the most widely used formula, with embedded probabilities, in human history".

This widespread use and practicality is extended by further research. In particular, stochastic volatility models have been introduced to address the volatility smiles observed in option markets and the heavy tails and high peaks of the underlying asset distributions. Hull and White [19], Chesney and Scott [2], Stein and Stein [29], Heston [17], and Hagan *et al.* [16] among many others assume a bivariate diffusion framework in which a separate stochastic process represents the dynamics of asset price volatility. In all these models, the asset price process S_t and its volatility factor process Y_t satisfy the following stochastic differential equations

$$\begin{aligned} dS_t &= S_t\left[\mu dt + f(Y_t)dW_t\right], \\ dY_t &= \mu^Y(Y_t)dt + \sigma^Y(Y_t)dZ_t, \end{aligned}$$

with correlated Brownian motions W, Z and different choices for the functions $\mu^Y(y)$, $\sigma^Y(y)$ and $f(y)$.

In this paper, we construct an approach that provides a *recombining* tree approximation for all stochastic volatility models of the above type. This approximation as the Cox, Ross & Rubinstein (CRR) model easily constructs a discrete time financial market that itself is arbitrage free and as such allows for simple analysis of related complex instruments.

For specificity, we implement our methodology on the Heston model. As well known, among stochastic volatility models it deserves special attention because of its ability to provide closed form solutions for European options through Fourier transform. This unique feature allows for an efficient and quick calibration of the model to European options. However, for derivative products with early exercise features such as American options, closed-form solutions do not exist even under the Heston model. Hence, numerical methods such as binomial tree, finite difference schemes or Monte Carlo simulation have to be used to evaluate American and other exotic options under stochastic volatility models.

In any market with a non-constant volatility, the CRR methodology encounters a basic difficulty. Indeed, since the volatility changes at each time, the nodes do not recombine on the lattice and this fact results in an exponential and thus a computationally explosive tree that cannot be used in many realistic situations. Nelson and Ramaswamy [24] were the first to construct a computationally simple binomial process which approximates a diffusion process given in the form

$$dY_t = \mu(Y, t)dt + \sigma(Y, t)dZ_t.$$

They solve the node recombination problem by transforming the process given in the above equation into a process $X(Y,t)$ such that the instantaneous volatility of the transformed process is constant. Hilliard and Schwartz [18] follow this method to develop binomial trees for continuous-time risk-neutralized diffusion processes of a special form.

Our main tool is to apply correlated random walks in order to approximate diffusion processes. A correlated random walk is a generalized random walk in the sense that the increments are not identically and independently distributed, but they only satisfy some Markov type of conditions. The exact definition will be given in Section 2.3. These processes naturally lie on a grid, and their Markov structure allows for an efficient computation of option prices.

The idea to use correlated random walks for approximating diffusion processes goes back to Gruber and Schweizer [15] and to Kusuoka [22]. In [15], the authors prove a convergence result for one dimensional diffusion process that satisfies stronger regularity conditions than that appear in stochastic volatility models. In [22], Kusuoka uses (also in one dimension) an original technique to modify random walks in order to get a diffusion in the limit. Again the regularity conditions that he assumes are stronger than that are required in stochastic volatility models.

Our approach is also similar to that of Kusuoka and modifies the correlated random walks on a multi-dimensional Binomial tree by adding a predictable process times $\sqrt{h}$ where h is the size of the time step. We then use this freedom to choose the predicable process together with an appropriate choice of the conditional probabilities to construct a Markov process that weakly converges to the stochastic volatility model. This construction is explained in Section 2.3. The weak convergence of our approximation is given in Section 2.4. Then, the approximating martingale measures are constructed so that the modified tree under these measures asymptotically matches the first two conditional moments. This fact allows for a straightforward convergence proof. We also note that this approach was successively used by the last two authors [9] to prove convergence of a market with trading costs.

Our extensive numerical experimentation is reported in our final section. In general, weak convergence does not provide any error estimation. However, binomial type approximations of diffusion models have a convergence rate of $\sqrt{\Delta t}$ which we accept it to be true.

We leave the detailed description of the computational studies to that section and here simply state that our algorithm works efficiently compared to all existing methods for the Heston model.

We emphasize that our tool can also be applied for a general type of stochastic volatility models (see Remark 2.4.2). There is also GARCH approach to stochastic volatility models that we refer the reader to Duan [10, 11, 12, 13], Nelson [23], Ritchken & Trevor [27], and the references therein.

Clearly, there are several other successful computational approaches to stochastic models, including the ones based on partial differential equations, semi-analytic methods and Monte-Carlo simulations. Here we do not survey all these results but compare our numerical results with the appropriate ones in the section that outlines our numerical experimentations.

In the literature, tree based methods have also been considered. Beliaeva & Nawalkha [2] is the most recent of these studies and we refer the reader to [2] and the references therein for these studies. However, our approach differs from these earlier studies in two fundamental points. Firstly, our approximation is recombinant by construction while in the previous studies recombination is achieved through truncation. Also, our tree is arbitrage free and we provide a proof of convergence.

2.2 The Heston Model

Consider the Heston model,

$$\begin{aligned} dS_t &= S_t(rdt + \sqrt{\nu}_t dW_t), \\ d\nu_t &= \kappa(\theta - \nu_t)dt + \eta\sqrt{\nu}_t d\widetilde{W}_t, \end{aligned}$$

with initial conditions $S_0, \nu_0 > 0$, given positive parameters r, κ, θ, η and two Brownian motions $W, \widetilde{W}$ with a constant correlation $\rho \in (-1, 1)$. The constant $r > 0$ is the interest rate and S is the stock price process. As it is standard, we also assume that

$$2\kappa\theta > \eta^2.$$

Then, the Heston equation has a unique *positive* solution in $\mathbb{R}^2_+$, see for instance [7].

The main goal of this paper is to construct a discrete approximation of this model. For this purpose, it is more convenient to work with a transformed system of affine equations driven by independent Brownian motions. Therefore, we set

$$x_t := \ln S_t, \qquad y_t := \frac{\nu_t}{\eta} - \rho x_t,$$

so that

$$\begin{aligned} dx_t &= \mu_x(x_t, y_t)dt + \sqrt{\eta}\,\sigma(x_t, y_t)dW_t, \\ dy_t &= \mu_y(x_t, y_t)dt + \sqrt{\eta(1-\rho^2)}\,\sigma(x_t, y_t)dB_t, \end{aligned} \tag{2.2.1}$$

where

$$\begin{aligned} &\mu_x(x,y) := r - \frac{1}{2}\eta(y+\rho x), \quad \mu_y(x,y) := \frac{\kappa\theta}{\eta} - \rho r + \frac{1}{2}\left(\rho\eta - 2\kappa\right)(y+\rho x), \\ &B_t := \frac{W_t - \rho\widetilde{W}_t}{\sqrt{1-\rho^2}}, \qquad \sigma(x,y) := \sqrt{(y+\rho x)^+}, \end{aligned}$$

and $z^+ = \max(0, z)$. One may directly verify that B is also a standard Brownian motion independent of W.

2.3 Derivation of the approximation

We fix a time horizon, or equivalently a maturity, $T > 0$ and a time discretization

$$h := \frac{T}{n},$$

with a large integer n. We then use two dimensional correlated random walks to approximate the diffusion processes given by (2.2.1). Indeed, consider the random walks

$\{X_k^{(n)}, Y_k^{(n)}\}_{k=0}^{n}$ of the form

$$X_k^{(n)} := x_0 + \sqrt{h\eta}\ \sum_{i=1}^{k} \xi_i^X, \tag{2.3.1}$$

$$Y_k^{(n)} := y_0 + \sqrt{h\eta(1-\rho^2)}\ \sum_{i=1}^{k} \xi_i^Y, \tag{2.3.2}$$

where $x_0 := \ln(s_0)$, $y_0 := (\nu_0/\eta) - \rho x_0$ and (ξ^X, ξ^Y)'s are random variables with values in $\{-1, 1\}$. In the sequel, we always use the initial data

$$\xi_0^X = \xi_0^Y = 0.$$

We construct a probabilistic structure so that the four tuple $(X_k^{(n)}, Y_k^{(n)}, \xi_k^X, \xi_k^Y)$ forms a Markov chain weakly approximating the solution of (2.2.1). To achieve this we also need to introduce a modification of this discrete Markov chain. Indeed, for given *predictable* processes $\hat{\alpha}, \hat{\beta}$, we introduce

$$\hat{X}_k^{(n)} := X_k^{(n)} + \sqrt{h\eta}\ \hat{\alpha}_k\ \xi_k^X, \tag{2.3.3}$$

$$\hat{Y}_k^{(n)} := Y_k^{(n)} + \sqrt{h\eta(1-\rho^2)}\ \hat{\beta}_k\ \xi_k^Y, \qquad k = 1, \ldots, n. \tag{2.3.4}$$

Clearly, the convergence of (X, Y) is equivalent to that of $(\hat{X}, \hat{Y})$ as

$$\|\hat{X}^{(n)} - X^{(n)}\| = O(\sqrt{h}), \qquad \|\hat{Y}^{(n)} - Y^{(n)}\| = O(\sqrt{h}),$$

where for any exponent k, we use the standard notation $O(h^k)$ to denote a generic random variable of the order h^k and $o(h^k)$ denotes a random variable that converges to zero after divided by h^k.

Our goal is to construct a sequence of probability measures $\mathbb{P}^{(n)}$ and stochastic processes $\hat{\alpha}^{(n)}, \hat{\beta}^{(n)}$ such that

$$\{(\hat{X}_{[nt/T]}^{(n)}, \hat{Y}_{[nt/T]}^{(n)})\}_{t=0}^{T} \Rightarrow \{(x_t, y_t)\}_{t=0}^{T},$$

where $\Rightarrow$ denotes weak convergence. We provide the definitions in the next section.

In view of the martingale convergence Theorem 7.4.1 in [14], to establish this convergence, it is essentially sufficient to match the first and the second conditional moments. Indeed, for a positive integer k, set

$$\mathcal{F}_k = \sigma\{\xi_1^X, ..., \xi_k^X, \xi_1^Y, ..., \xi_k^Y\},$$

and let $\mathbb{E}_k^{(n)}[\,\cdot\,]$ be the conditional expectation $\mathbb{E}^{(n)}[\,\cdot\,|\mathcal{F}_k]$ with respect to the probability measure $\mathbb{P}^{(n)}$. Then, the moment matching conditions are the following equations,

$$\mathbb{E}_{k-1}^{(n)}[\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)}] = \mu_x(X_{k-1}^{(n)}, Y_{k-1}^{(n)})h + o(h), \tag{2.3.5}$$

$$\mathbb{E}_{k-1}^{(n)}[\hat{Y}_k^{(n)} - \hat{Y}_{k-1}^{(n)}] = \mu_y(X_{k-1}^{(n)}, Y_{k-1}^{(n)})h + o(h), \tag{2.3.6}$$

$$\mathbb{E}_{k-1}^{(n)}\left[(\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)})^2\right] = \eta\sigma^2(X_{k-1}^{(n)}, Y_{k-1}^{(n)})h + o(h), \tag{2.3.7}$$

$$\mathbb{E}_{k-1}^{(n)}\left[(\hat{Y}_k^{(n)} - \hat{Y}_{k-1}^{(n)})^2\right] = \eta(1-\rho^2)\sigma^2(X_{k-1}^{(n)}, Y_{k-1}^{(n)})h + o(h). \tag{2.3.8}$$

We also need conditions on the the covariances. However, since W and B in (2.2.1) are independent, this condition is simply reduced to the requirement that ξ_k^X and ξ_k^Y are conditionally independent given $\mathcal{F}_{k-1}$.

Observe that we need to solve four equations and the number of unknowns or parameters to choose are four as well; the corrections $\hat{\alpha}, \hat{\beta}$ and two probabilities,

$$p_k := \mathbb{P}_{k-1}^{(n)}(\xi_k^X = 1), \qquad q_k := \mathbb{P}_{k-1}^{(n)}(\xi_k^Y = 1). \tag{2.3.9}$$

This construction would provide a financial market which is asymptotically arbitrage free. However, a slight modification of the above procedure would also ensure that each discrete market itself is free of arbitrage. In our model, the discrete stochastic process

$$\{\exp(-rkh)\exp(\hat{X}_k^{(n)})\}_{k=0}^{n},$$

is the approximation of the discounted price process. Hence, we replace the first order condition (2.3.5) by requiring that above process is a martingale, i.e.,

$$\mathbb{E}_{k-1}^{(n)}[\exp(-rh)\exp(\hat{X}_k^{(n)}) - \exp(\hat{X}_{k-1}^{(n)})] = 0. \tag{2.3.10}$$

In fact, (2.3.5) and (2.3.10) are asymptotically equivalent and both would be sufficient to prove convergence. However, in our numerical experimentation we observe that this modification is substantially better than the non-modified version. We continue by constructing $\mathbb{P}^{(n)}$ and $\hat{\alpha}^{(n)}, \hat{\beta}^{(n)}$ satisfying the equations (2.3.10), and (2.3.6)–(2.3.8). Indeed, by (2.3.10) we directly calculate that

$$(1+\hat{\alpha}_k)\mathbb{E}_{k-1}^{(n)}[\xi_k^X] - \hat{\alpha}_{k-1}\xi_{k-1}^X = o(h).$$

Hence,

$$(1+\hat{\alpha}_k^{(n)})(\hat{\alpha}_{k-1})\mathbb{E}_{k-1}^{(n)}[\xi_k^X]\xi_{k-1}^X = (\hat{\alpha}_{k-1})^2 + o(h).$$

We use this and calculate that

$$\mathbb{E}_{k-1}^{(n)}\left((\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)})^2\right) = \eta h\left((1+\hat{\alpha}_k)^2 - (\hat{\alpha}_{k-1}^{(n)})^2 + o(h)\right) + o(h).$$

We expect that the difference $\hat{\alpha}_k - \hat{\alpha}_{k-1}$ to be of order h. Hence, the above expression simplifies to

$$\mathbb{E}_{k-1}^{(n)}\left((\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)})^2\right) = \eta h\,(1+2\hat{\alpha}_k) + o(h).$$

We now compare the above equation with (2.3.7) to conclude that

$$1+2\hat{\alpha}_k = \sigma^2(X_{k-1}^{(n)}, Y_{k-1}^{(n)}) + o(h).$$

Using (2.3.6) and (2.3.8), we obtain the same equation for $\hat{\beta}$. Hence, we conclude that

$$\hat{\alpha}_k = \hat{\beta}_k = \frac{\sigma^2(X_{k-1}^{(n)}, Y_{k-1}^{(n)}) - 1}{2} + o(h).$$

We use the above identity and the freedom on the order $o(h)$ to define the processes $\hat{\alpha}, \hat{\beta}$ below. The below definition contains a certain truncation that is within the $o(h)$ margin. Although this correction is asymptotically small, it allows us to obtain several bounds in the convergence proof and also enables to construct transition probabilities that always remain in the unit interval (see (2.3.12), below). So we now define

$$\hat{\alpha}_k := \hat{\beta}_k := \frac{\max\left\{A_n, \sigma^2(X^{(n)}_{k-1}, Y^{(n)}_{k-1})\right\} - 1}{2}, \quad 1 \le k \le n, \tag{2.3.11}$$

where

$$A_n = \left(\frac{\kappa\theta}{\eta} + |\rho| r\right) \sqrt{\frac{h}{\eta(1-\rho^2)}},$$

and we set

$$\hat{\alpha}^{(n)}_0 = \hat{\beta}^{(n)}_0 = 0.$$

To reiterate once again the function A_n is chosen to ensure that the probabilities that are defined in (2.3.12), below, remain in the unit interval. Although, this is clearly crucial for our analysis, in our numerical implementation we do not use this truncation and instead modify (2.3.12) to ensure that these are true probabilities.

The above construction together with the conditional independence of the increments ensures the second moment matching. We now use the first order conditions (2.3.10) and (2.3.6) to construct the transition probabilities. Indeed, recall that by (2.3.9),

$$p_k := \mathbb{P}^{(n)}_{k-1}(\xi^X_k = 1),$$

and rewrite (2.3.10) as

$$\begin{aligned} p_k \exp\left(\sqrt{h\eta}\ [(1+\hat{\alpha}_k) - \hat{\alpha}_{k-1}\xi^X_{k-1}]\right) \\ +(1-p_k)\exp\left(-\sqrt{h\eta}\ [(1+\hat{\alpha}_k) + \hat{\alpha}_{k-1}\xi^X_{k-1}]\right) = \exp(rh). \end{aligned}$$

This implies that p_k must be given by

$$p_k = \frac{\exp(rh + \sqrt{\eta h}\ \hat{\alpha}_{k-1}\xi^X_{k-1}) - \exp(-\sqrt{\eta h}\ (1+\hat{\alpha}_k))}{\exp(\sqrt{\eta h}\ (1+\hat{\alpha}_k)) - \exp(-\sqrt{\eta h}\ (1+\hat{\alpha}_k))}. \tag{2.3.12}$$

In view of the truncation introduced in (2.3.11), $p_k \in [0,1]$ for all large n.

We now recall that

$$q_k := \mathbb{P}^{(n)}_{k-1}(\xi^Y_k = 1),$$

and use (2.3.6) to arrive at

$$q_k = \frac{1}{2} + \frac{\hat{\alpha}_{k-1}}{2(1+\hat{\alpha}_k)}\xi^Y_{k-1} + \frac{\sqrt{h}\ \mu_y(X^{(n)}_{k-1}, Y^{(n)}_{k-1})}{2\sqrt{\eta(1-\rho^2)}\ (1+\hat{\alpha}_k)}.$$

Since q_k must take values in the unit interval, we modify it in the following way,

$$q_k = \left(\min\left\{1, \frac{1}{2} + \frac{\hat{\alpha}_{k-1}}{2(1+\hat{\alpha}_k)}\xi^Y_{k-1} + \frac{\sqrt{h}\ \mu_y(X^{(n)}_{k-1}, Y^{(n)}_{k-1})}{2\sqrt{\eta(1-\rho^2)}(1+\hat{\alpha}_k)}\right\}\right)^+. \tag{2.3.13}$$

Set

$$\Xi_k \quad := \quad (X^{(n)}_k, Y^{(n)}_k, \xi^X_k, \xi^Y_k).$$

Then, we claim that Ξ is a Markov process. Indeed, recall that the independence of the Brownian motions in (2.2.1) implies the conditional independence of the increments ξ^X and ξ^Y. Hence,

$$\mathbb{P}^{(n)}\left(\xi^X_k = a, \xi^Y_k = b \mid \Xi_{k-1}\right) = \mathbb{P}^{(n)}_{k-1}(\xi^X_k = a)\ \mathbb{P}^{(n)}_{k-1}(\xi^Y_k = b). \tag{2.3.14}$$

Moreover, in view of (2.3.1) and (2.3.2), the set

$$\{X^{(n)}_k = X^{(n)}_{k-1} + c,\ Y^{(n)}_k = Y^{(n)}_{k-1} + d,\ \xi^X_k = a,\ \xi^Y_k = b\}$$

is empty unless $c = a\eta h$ and $d = b\eta h\sqrt{1-\rho^2}$, and in this case it is equal to $\{\xi^X_k = a,\ \xi^Y_k = b\}$. Therefore, the transition probabilities of the process Ξ are determined by

$$\mathbb{P}^{(n)}\left(\xi^X_k = 1, \xi^Y_k = 1 \mid \Xi_{k-1}\right) = p_k\ q_k.$$

Moreover, there is a simple transformation between Ξ_k and

$$\hat{\Xi}_k := (\hat{X}^{(n)}_k, \hat{Y}^{(n)}_k, \xi^X_k, \xi^Y_k).$$

Hence, one may consider the process $\hat{\Xi}$ as the basic approximating Markov process.

2.4 Main convergence result

In this section, we first briefly recall the concept of weak convergence of probability measures and then state our main convergence result. For more information on weak convergence, we refer the reader to the books of Billingsley [3] and Ethier & Kurtz [14].

For any *càdlàg* stochastic process $\{Z(t)\}_{t=0}^{T}$ with values in some Euclidean space $\mathbb{R}^d$, let $\mathbb{P}^Z$ be the distribution of Z on the canonical space $\mathbb{D}([0,T];\mathbb{R}^d)$ equipped with the Skorohod topology (for details see [3]), i.e. for any Borel set $D \subset \mathbb{D}([0,T];\mathbb{R}^d)$, $\mathbb{P}^Z(D) = \mathbb{P}\{Z \in D\}$. For a sequence of $\mathbb{R}^d$-valued, stochastic processes $Z^{(n)}$ we use the notation $Z^{(n)} \Rightarrow Z$ to indicate that the probability measures $\mathbb{P}^{Z^{(n)}}$, converge vaguely to $\mathbb{P}^Z$ on the space $\mathbb{D}([0,T];\mathbb{R}^d)$.

We are now ready to state the main convergence theorem which is the main theoretical foundation of our numerical scheme. It will be proved in Section 2.6.

Theorem 2.4.1. For any $n \in \mathbb{N}$, let $\mathbb{P}^{(n)}$ be the probability measure defined by (2.3.14). Consider the stochastic processes $\{X^{(n)}_{[nt/T]}\}_{t=0}^{T}$, $\{\hat{X}^{(n)}_{[nt/T]}\}_{t=0}^{T}$ and $\{Y^{(n)}_{[nt/T]}\}_{t=0}^{T}$ under $\mathbb{P}^{(n)}$. Let (x,y) be the unique solution of (2.2.1). Then,

$$\{(X^{(n)}_{[nt/T]}, Y^{(n)}_{[nt/T]})\}_{t=0}^{T} \Rightarrow \{(x_t, y_t)\}_{t=0}^{T} \tag{2.4.1}$$

and

$$\{(\hat{X}^{(n)}_{[nt/T]}, Y^{(n)}_{[nt/T]})\}_{t=0}^{T} \Rightarrow \{(x_t, y_t)\}_{t=0}^{T} \tag{2.4.2}$$

on the space $\mathbb{D}([0,T]) \times \mathbb{D}([0,T])$.

Remark 2.4.2. For the Heston model, one applies a transformation that decorrelates the Brownian motions. However, this decorrelation is not necessary and used only to simplify the procedure. Indeed, consider a general two dimensional diffusion

$$\begin{aligned} dx_t &= \mu_x(x_t, y_t)dt + \sigma_x(x_t, y_t)dW_t, \\ dy_t &= \mu_y(x_t, y_t)dt + \sigma_y(x_t, y_t)d\tilde{W}_t, \end{aligned}$$

where $W, \widetilde{W}$ are two standard Brownian motions with a correlation ρ. Introduce the two dimensional correlated random walk $\{X_k^{(n)}, Y_k^{(n)}\}_{k=0}^{n}$ by

$$\begin{aligned} X_k^{(n)} &:= x_0 + \sqrt{h}\sum_{i=1}^{k} \xi_i^X, \\ Y_k^{(n)} &:= y_0 + \sqrt{h}\sum_{i=1}^{k} \xi_i^Y. \end{aligned}$$

As before, we consider a small modification of the correlated random walks

$$\begin{aligned} \hat{X}_k^{(n)} &:= X_k^{(n)} + \sqrt{h}\ \hat{\alpha}_k\ \xi_k^X, \\ \hat{Y}_k^{(n)} &:= Y_k^{(n)} + \sqrt{h}\hat{\beta}_k \xi_k^Y, \qquad k = 1, \ldots, n. \end{aligned}$$

In this case, the moment matching conditions are the following equations,

$$\begin{aligned} &\mathbb{E}_{k-1}^{(n)}[\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)}] = \mu_x(X_{k-1}^{(n)}, Y_{k-1}^{(n)})h + o(h), \\ &\mathbb{E}_{k-1}^{(n)}[\hat{Y}_k^{(n)} - \hat{Y}_{k-1}^{(n)}] = \mu_y(X_{k-1}^{(n)}, Y_{k-1}^{(n)})h + o(h), \\ &\mathbb{E}_{k-1}^{(n)}\left[(\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)})^2\right] = \sigma_x^2(X_{k-1}^{(n)}, Y_{k-1}^{(n)})h + o(h), \\ &\mathbb{E}_{k-1}^{(n)}\left[(\hat{Y}_k^{(n)} - \hat{Y}_{k-1}^{(n)})^2\right] = \sigma_y^2(X_{k-1}^{(n)}, Y_{k-1}^{(n)})h + o(h), \\ &\mathbb{E}_{k-1}^{(n)}\left[(\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)})(\hat{Y}_k^{(n)} - \hat{Y}_{k-1}^{(n)})\right] = \sigma_x(X_{k-1}^{(n)}, Y_{k-1}^{(n)})\sigma_y(X_{k-1}^{(n)}, Y_{k-1}^{(n)})\rho h + o(h). \end{aligned}$$

We solve these equations as in the Heston case and obtain that

$$\hat{\alpha}_k = \frac{\sigma_x^2(X_{k-1}^{(n)}, Y_{k-1}^{(n)}) - 1}{2}, \quad \hat{\beta}_k = \frac{\sigma_y^2(X_{k-1}^{(n)}, Y_{k-1}^{(n)}) - 1}{2}.$$

The transition probabilities are also given by,

$$\begin{aligned} \mathbb{P}_{k-1}^{(n)}(\xi_k^X = 1, \xi_k^Y = 1) &= \frac{1}{4} + \frac{\hat{\alpha}_{k-1}\xi_{k-1}^X + \mu_x\sqrt{h}}{4(1+\hat{\alpha}_k)} + \frac{\hat{\beta}_{k-1}\xi_{k-1}^Y + \mu_y\sqrt{h}}{4(1+\hat{\beta}_k)} \\ &\quad + \frac{\rho\sigma_x\sigma_y + \hat{\alpha}_{k-1}\hat{\beta}_{k-1}\xi_{k-1}^X\xi_{k-1}^Y}{4(1+\hat{\alpha}_k)(1+\hat{\beta}_k)}, \end{aligned}$$

$$\mathbb{P}^{(n)}_{k-1}(\xi^X_k = 1, \xi^Y_k = -1) = \frac{1}{4} + \frac{\hat{\alpha}_{k-1}\xi^X_{k-1} + \mu_x\sqrt{h}}{4(1+\hat{\alpha}_k)} - \frac{\hat{\beta}_{k-1}\xi^Y_{k-1} + \mu_y\sqrt{h}}{4(1+\hat{\beta}_k)} - \frac{\rho\sigma_x\sigma_y + \hat{\alpha}_{k-1}\hat{\beta}_{k-1}\xi^X_{k-1}\xi^Y_{k-1}}{4(1+\hat{\alpha}_k)(1+\hat{\beta}_k)},$$

$$\mathbb{P}^{(n)}_{k-1}(\xi^X_k = -1, \xi^Y_k = 1) = \frac{1}{4} - \frac{\hat{\alpha}_{k-1}\xi^X_{k-1} + \mu_x\sqrt{h}}{4(1+\hat{\alpha}_k)} + \frac{\hat{\beta}_{k-1}\xi^Y_{k-1} + \mu_y\sqrt{h}}{4(1+\hat{\beta}_k)} - \frac{\rho\sigma_x\sigma_y + \hat{\alpha}_{k-1}\hat{\beta}_{k-1}\xi^X_{k-1}\xi^Y_{k-1}}{4(1+\hat{\alpha}_k)(1+\hat{\beta}_k)},$$

$$\mathbb{P}^{(n)}_{k-1}(\xi^X_k = -1, \xi^Y_k = -1) = \frac{1}{4} - \frac{\hat{\alpha}_{k-1}\xi^X_{k-1} + \mu_x\sqrt{h}}{4(1+\hat{\alpha}_k)} - \frac{\hat{\beta}_{k-1}\xi^Y_{k-1} + \mu_y\sqrt{h}}{4(1+\hat{\beta}_k)} + \frac{\rho\sigma_x\sigma_y + \hat{\alpha}_{k-1}\hat{\beta}_{k-1}\xi^X_{k-1}\xi^Y_{k-1}}{4(1+\hat{\alpha}_k)(1+\hat{\beta}_k)},$$

where in the above formulae, functions $\mu_x, \mu_y, \sigma_x, \sigma_y$ are all evaluated at $(X^{(n)}_{k-1}, Y^{(n)}_{k-1})$. However, the above terms do not necessarily lie in the interval $[0, 1]$. In that case, we apply a truncation of the form $\min(1, \max(0, \cdot))$. $\square$

Remark 2.4.3. We emphasize that our approximation method using correlated random walks and the above convergence result can easily be extended to more general multidimensional diffusions. The key idea is the introduction of $\hat{X}$ type processes which differ from the original random walk X only by a predictable process $\hat{\alpha}$ times the increment ξ^X. We then use this freedom (namely, the function $\hat{\alpha}$) to construct transition probabilities that match the first and the second conditional moments of the original diffusion. The approximating process has the essentially same dimension as the original diffusion process. However, we need to augment the state space by adding the increments like ξ^X. But these increments take values in the discrete set $\{-1, +1\}$ so do not increase the complexity of the approximation. $\square$

Our next remark is towards American options.

Remark 2.4.4. In general, the usual weak convergence is not sufficient for the convergence of American options prices. Indeed, the latter also requires the "good" behavior of the filtrations. In his unpublished manuscript (see [1], Sections 15–16), David Aldous introduced the concept of extended weak convergence to address this problem. Briefly his definition is as follows. A sequence $Z^{(n)} : \Omega_n \to \mathbb{D}([0,T];\mathbb{R}^d)$, *extended weak converges* to a stochastic process $Z : \Omega \to \mathbb{D}([0,T];\mathbb{R}^d)$, if for any k and continuous bounded functions $\psi_1, ..., \psi_k \in C(\mathbb{D}([0,T];\mathbb{R}^d))$,

$$(Z^{(n)}, Z^{n,1}, ..., Z^{n,k}) \Rightarrow (Z, Z^{(1)}, ..., Z^{(k)}) \text{ in } \mathbb{D}([0,T];\mathbb{R}^{d+k}),$$

where for any $t \leq T$, $1 \leq i \leq k$ and $n \in \mathbb{N}$,

$$Z^{n,i}_t = \hat{E}^{(n)}(\psi_i(Z^{(n)}) \mid \mathcal{F}^{Z^{(n)}}_t), \qquad Z^{(i)}_t = \hat{E}(\psi_i(Z) \mid \mathcal{F}^Z_t),$$

$\hat{E}^{(n)}$ denotes the expectation on the probability space on which $Z^{(n)}$ is defined and $\hat{E}$ denotes the expectation on the probability space on which Z is defined. In the formulas above $\mathcal{F}^{Z^{(n)}}$ and $\mathcal{F}^Z$ are the filtrations which are generated by $Z^{(n)}$ and Z, respectively. The notion of extended weak convergence provides (in addition to the standard weak convergence of stochastic processes) convergence of filtrations. In particular, Aldous proved (see [1] Section 17) that under uniform integrability of the payoffs, extended weak convergence implies convergence of optimal stopping values. However, it is known that when the proof of weak convergence relies on martingale techniques (like our proof), then the standard weak convergence implies extended weak convergence. For details, we refer the reader to [1], Section 21.

2.5 Discrete pricing equations

In this section, we apply the approximation developed in Section 2.3 to price American put and lookback options.

2.5.1 American Put

Consider an American Put option with a strike price K. We are interested in approximating its value given by,

$$V = \sup_{\tau \in \mathcal{T}_{[0,T]}} \mathbb{E}\left(e^{-r\tau}(K - S_\tau)^+\right),$$

where $\mathcal{T}_{[0,T]}$ is the set of all stopping times with respect to the filtration generated by S, with values in the set $[0,T]$. We approximate the discounted stock price by the discrete time martingales

$$\{e^{-rkh}e^{\hat{X}_k^{(n)}}\}_{k=0}^{n}, \quad n \in \mathbb{N},$$

constructed in Section (2.3). For any $n \in \mathbb{N}$, let $\mathcal{T}_n$ be the set of all stopping times with respect to the filtration $\mathcal{F}_k$ (again constructed in Section (2.3)), with values in the set $\{0,1,...,n\}$. Define,

$$V^{(n)} := \max_{\tau \in \mathcal{T}_n} \mathbb{E}^{(n)}\left(e^{-r\tau h}(K - S_0 e^{\hat{X}_\tau^{(n)}})^+\right).$$

In view of Theorem (2.4.1) and Remark (2.4.4), we directly conclude that

$$\lim_{n\to\infty} V^{(n)} = V.$$

Next, we describe a dynamical programming algorithm for the calculation of $V^{(n)}$. Observe that for a given $k \in \{0,\dots,n\}$ the random variables $X_k^{(n)}$ and $Y_k^{(n)}$ take values on the grid

$$\begin{aligned} x_0 + (2l-k)\sqrt{\eta h}, &\qquad 0 \le l \le k, \\ y_0 + (2m-k)\sqrt{\eta(1-\rho^2)h}, &\qquad 0 \le m \le k, \end{aligned}$$

respectively. For non-negative integers $m, l \le k \le n$ and $\xi_x, \xi_y \in \{-1,+1\}$, let

$$V_k^{(n)}(l,m,\xi_x,\xi_y)$$

be the value of the option at time k when the Markov process is given by

$$\begin{aligned} \Xi_k &= (X_k^{(n)},\ Y_k^{(n)},\ \xi_k^X,\ \xi_k^Y) = F_k(l,m,\xi_x,\xi_y) \\ &:= (x_0 + (2l-k)\sqrt{\eta h},\ y_0 + (2m-k)\sqrt{\eta(1-\rho^2)h},\ \xi_x,\ \xi_y). \end{aligned}$$

The above function F_k is invertible with an inverse F_k^{-1}. We sometimes, with an abuse of notation, write

$$V_{k-1}^{(n)}(\Xi) = V^{(n)}(F_{k-1}^{-1}(\Xi)),$$

for any four tuple Ξ given by $F_{k-1}(l, m, \xi_x, \xi_y)$ for some (l, m, ξ_x, ξ_y). With this convention, it is not straightforward to state the dynamic programming equation (see, for instance [26], Chapter 1),

$$V_{k-1}^{(n)}(\Xi) = \max\left\{ \left(K - S_0 \exp(\hat{X}_{k-1})\right)^+ \ , \ \mathbb{E}^{(n)} \left[V_k^{(n)}(\Xi_k) \ \middle| \ \Xi_{k-1} = \Xi \ \right] \right\}. \qquad (2.5.1)$$

We continue by rewriting the dynamic programming equation in an algorithmic manner. In view of ((2.3.11))–((2.3.13)), for any $1 \leq k \leq n$ and $0 \leq l, m \leq k-1$, we define,

$$\begin{aligned} \mathcal{X}_k &:= x_0 + (2l - k)\sqrt{\eta h}, \\ \mathcal{Y}_k &:= y_0 + (2m - k)\sqrt{\eta(1-\rho^2)h}, \end{aligned}$$

where both of the above are functions of (l, m), but this dependence is suppressed in the notation. Similarly, we define two probabilities

$$\begin{aligned} p_k(l, m, \xi_x, \xi_y) &:= \frac{\exp(rh + \sqrt{\eta h}\ \Psi_{k-1}\ \xi_x) - \exp(-\sqrt{\eta h}\ \Psi_k)}{\exp(\sqrt{\eta h}\ \Psi_k) - \exp(-\sqrt{\eta h}\ \Psi_k)}, \\ q_k(l, m, \xi_x, \xi_y) &:= \left(\min\left\{ 1, \frac{1}{2} + \frac{\alpha_{k-1}(l - \xi_x, m - \xi_y)\ \xi_y}{2\ \Psi_k} + \frac{\sqrt{h}\ \mu_{y,k}}{2\ \sqrt{\eta(1-\rho^2)}\ \Psi_k} \right\} \right)^+, \end{aligned}$$

where $\alpha_0^{(n)} \equiv 0$ and

$$\begin{aligned} \alpha_k(l, m) &:= \frac{\max\left(A_n, \sigma^2\left(\mathcal{X}_{k-1}, \mathcal{Y}_{k-1}\right)\right) - 1}{2}, \\ \Psi_k &:= 1 + \alpha_k^{(n)}(l, m), \\ \mu_{y,k} &:= \mu_y(\mathcal{X}_{k-1}, \mathcal{Y}_{k-1}), \end{aligned}$$

As remarked earlier, in our actual numerical codes, we simply define $\alpha = (\sigma^2 - 1)/2$ without the truncation with A_n and instead truncate p_k, above, to ensure that it says within the unit interval.

Observe that

$$\begin{aligned} p_k(l,m,\xi_x,\xi_y) &= \mathbb{P}^{(n)}\left(\xi_k^X = 1 \mid \Xi_{k-1} = F_{k-1}(l,m,\xi_x,\xi_y)\ \right), \\ q_k(l,m,\xi_x,\xi_y) &= \mathbb{P}^{(n)}\left(\xi_k^Y = 1 \mid \Xi_{k-1} = F_{k-1}(l,m,\xi_x,\xi_y)\ \right). \end{aligned}$$

Moreover,

$$\mathbb{P}_{k-1}^{(n)}\left(\xi_k^X = 1\ ,\ \xi_k^Y = 1\ \right) = p_k(l,m,\xi_x,\xi_y)\ q_k(l,m,\xi_x,\xi_y).$$

One can easily obtain expressions for other three probabilities as well.

We are now ready to restate the dynamic programming equation (2.5.1). Indeed, $V_k^{(n)}(l,m,\xi_x,\xi_y)$ is the unique solution of the following recursive relations,

$$V_n^{(n)}(l,m,\xi_x,\xi_y) = \left(K - S_0 \exp\left(\sqrt{\eta h}\ \ (\mathcal{X}_n + \alpha_n \xi_x)\right)\right)^+,$$

and for $1 \le k \le n$,

$$V_{k-1}^{(n)}(l,m,\xi_x,\xi_y) = \max\left\{\left(K - S_0 \exp\left(\sqrt{\eta h}\ \ (\mathcal{X}_{k-1} + \alpha_{k-1}\xi_x)\right)\right)^+\ ,\ \mathcal{E}(V_k^{(n)})\ \right\},$$

where

$$\begin{aligned} \mathcal{E}(V_k^{(n)}) &= \mathbb{E}^{(n)}\left[V_k^{(n)}(\Xi_k)\ \Big|\ \Xi_{k-1} = F_{k-1}(l,m,\xi_x,\xi_y)\ \right], \\ &= \sum_{i,j=0}^{1} \mathbb{P}_{k-1}^{(n)}\left(\xi_k^X = 2i-1, \xi_k^Y = 2j-1\right) V_k^{(n)}(l+i,m+j,2i-1,2j-1), \\ &= \sum_{i,j=0}^{1} [1-i+(2i-1)p_k(l,m,\xi_x,\xi_y)]\ [1-j+(2j-1)q_k(l,m,\xi_x,\xi_y)] \\ &\qquad \times V_k^{(n)}(l+i,m+j,2i-1,2j-1). \end{aligned}$$

Then, our approximation is simply given by,

$$V_n = V_0^{(n)}(0,0,0,0).$$

2.5.2 Lookback options

Consider a lookback put option with a fixed strike K, i.e., an option with payoff $(K - \min_{0\leq t\leq T} S_t)^+$. Again, we want to approximate the price

$$\hat{V} = \mathbb{E}\left(e^{-rT}\left(K - \min_{0\leq t\leq T} S_t\right)^+\right).$$

Since the running minimum of the processes

$$\{\exp(X_k^{(n)})\}_{k=0}^{n}, \quad n \in \mathbb{N},$$

lies on a grid, we will use these processes instead of the martingale $\exp(\hat{X}_k^{(n)})$. The advantage of the processes $\exp(X_k^{(n)})$ becomes clear, when we describe the dynamical programming algorithm, below.

We set,

$$\hat{V}^{(n)} = \mathbb{E}^{(n)}\left(e^{-rT}\left(K - S_0 \exp\big(\min_{0\leq i\leq n} X_i^{(n)}\big)\right)^+\right). \tag{2.5.2}$$

By Theorem (2.4.1) we conclude that $\hat{V}^{(n)}$ converges to $\hat{V}$.

First, we observe that the random variable

$$z_k := \min_{0\leq i\leq k} \sum_{j=1}^{i} \xi_j^X$$

takes values on the grid $\{-k, 1-k, ..., 0\}$.

Using the notations and the conventions of the previous subsection, for $0 \leq k \leq n$, we let $\hat{V}_k^{(n)}(l, m, z, \xi_x, \xi_y)$ be the option price at time k. The extra state variable z denotes the value of the running minimum z_k at time k. Then, $\hat{V}^{(n)}$ is the unique solution of

$$\hat{V}_n^{(n)}(l, m, z, \xi_x, \xi_y) = \left(K - S_0 \exp\left(-\sqrt{\eta h}\ z\right)\right)^+,$$

and for $1 \le k \le n$,

$$\hat{V}_{k-1}^{(n)}(l, m, z, \xi_x, \xi_y) = \max\left\{ \left(K - S_0 \exp\left(-\sqrt{\eta h}\ z\right)\right)^+,\ \hat{\mathcal{E}}(V_k^{(n)}) \right\},$$

where

$$\begin{aligned} \hat{\mathcal{E}}(V_k^{(n)}) &= \sum_{i,j=0}^{1} \mathbb{P}_{k-1}^{(n)}\left(\xi_k^X = 2i-1, \xi_k^Y = 2j-1\right) \\ &\quad \times \hat{V}_k^{(n)}\left(l+i, m+j, z+\chi_{\{i=0, z+2l=k-1\}}, 2i-1, 2j-1\right), \end{aligned}$$

and χ_Q is the characteristic set of Q. Finally,

$$\hat{V}_n = \hat{V}_0^{(n)}(0,0,0,0,0).$$

2.6 Proof of Theorem 2.1

In this section we provide a proof of Theorem (2.4.1). Our main tool is the martingale convergence result of Theorem 7.4.1 in [14].

In view of ((2.3.1))–((2.3.4)) and ((2.3.11)), we have the following inequality for all sufficiently large n,

$$\begin{aligned} |\hat{X}_k^{(n)}| &\ge |X_k^{(n)}| - \frac{1}{3}(|X_k^{(n)}| + |Y_k^{(n)}| + 1), \\ |\hat{Y}_k^{(n)}| &\ge |Y_k^{(n)}| - \frac{1}{3}(|X_k^{(n)}| + |Y_k^{(n)}| + 1). \end{aligned}$$

Therefore,

$$|X_k^{(n)}| + |Y_k^{(n)}| \le 3(|\hat{X}_k^{(n)}| + |\hat{Y}_k^{(n)}| + 1), \quad k = 0, 1, ..., n. \tag{2.6.1}$$

This together with ((2.3.3))–((2.3.4)) and ((2.3.11)) implies that there exists a constant $c > 0$ satisfying,

$$|X_k^{(n)} - \hat{X}_k^{(n)}| + |Y_k^{(n)} - \hat{Y}_k^{(n)}| \le \frac{c(1 + |\hat{X}_k^{(n)}| + |\hat{Y}_k^{(n)}|)}{\sqrt{n}}, \quad k = 0, 1, ..., n. \tag{2.6.2}$$

It is sufficient to establish that

$$\{(\hat{X}^{(n)}_{[nt/T]}, \hat{Y}^{(n)}_{[nt/T]})\}_{t=0}^{T} \Rightarrow \{(x_t, y_t)\}_{t=0}^{T}. \tag{2.6.3}$$

Indeed, from ((2.6.2)) it follows that

$$\hat{X}^{(n)}_k - \frac{c(1+|\hat{X}^{(n)}_k|+|\hat{Y}^{(n)}_k|)}{\sqrt{n}} \le X^{(n)}_k \le \hat{X}^{(n)}_k + \frac{c(1+|\hat{X}^{(n)}_k|+|\hat{Y}^{(n)}_k|)}{\sqrt{n}},$$

$$\hat{Y}^{(n)}_k - \frac{c(1+|\hat{X}^{(n)}_k|+|\hat{Y}^{(n)}_k|)}{\sqrt{n}} \le Y^{(n)}_k \le \hat{Y}^{(n)}_k + \frac{c(1+|\hat{X}^{(n)}_k|+|\hat{Y}^{(n)}_k|)}{\sqrt{n}}.$$

From ((2.6.3)) it follows that the sequences

$$\left\{\left(\hat{X}^{(n)}_{[nt/T]} - \frac{c(1+|\hat{X}^{(n)}_{[nt/T]}|+|\hat{Y}^{(n)}_{[nt/T]}|)}{\sqrt{n}}, \hat{Y}^{(n)}_{[nt/T]} - \frac{c(1+|\hat{X}^{(n)}_{[nt/T]}|+|\hat{Y}^{(n)}_{[nt/T]}|)}{\sqrt{n}}\right)\right\},$$

$$\left\{\left(\hat{X}^{(n)}_{[nt/T]} + \frac{c(1+|\hat{X}^{(n)}_{[nt/T]}|+|\hat{Y}^{(n)}_{[nt/T]}|)}{\sqrt{n}}, \hat{Y}^{(n)}_{[nt/T]} + \frac{c(1+|\hat{X}^{(n)}_{[nt/T]}|+|\hat{Y}^{(n)}_{[nt/T]}|)}{\sqrt{n}}\right)\right\}$$

converge weakly to $\{(x_t, y_t)\}_{t=0}^{T}$. Thus, Theorem (2.4.1) follows from ((2.6.3)). For any $0 \le k \le n$, set

$$\begin{aligned}
A^{n,x}_k &= \sum_{j=1}^{k} \mathbb{E}^{(n)}_{j-1}\left(\hat{X}^{(n)}_j - \hat{X}^{(n)}_{j-1}\right), \quad A^{n,y}_k = \sum_{j=1}^{k} \mathbb{E}^{(n)}_{j-1}\left(\hat{Y}^{(n)}_j - \hat{Y}^{(n)}_{j-1}\right),\\
M^{n,x}_k &= \hat{X}^{(n)}_k - A^{n,x}_k, \quad M^{n,y}_k = \hat{Y}^{(n)}_k - A^{n,y}_k,\\
A^{n,x,x}_k &= \sum_{j=1}^{k} \mathbb{E}^{(n)}_{j-1}\left((M^{n,x}_j - M^{n,x}_{j-1})^2\right), \quad A^{n,y,y}_k = \sum_{j=1}^{k} \mathbb{E}^{(n)}_{j-1}\left((M^{n,y}_j - M^{n,y}_{j-1})^2\right),\\
A^{n,x,y}_k &= \sum_{j=1}^{k} \mathbb{E}^{(n)}_{j-1}\left((M^{n,x}_j - M^{n,x}_{j-1})(M^{n,y}_j - M^{n,y}_{j-1})\right).
\end{aligned}$$

Notice that the processes $A^{n,x}, A^{n,y}, A^{n,x,x}, A^{n,y,y}, A^{n,x,y}$ are predictable and the processes $M^{n,x}, M^{n,y}$ are martingales.

We now fix a large $N > 0$ and define the stopping times by,

$$\sigma_n = \min\{k : |\hat{X}_k^{(n)}| + |\hat{Y}_k^{(n)}| \geq N\} \wedge n, \qquad n \in \mathbb{N}.$$

Using (2.3.1), (2.3.2) and (2.6.2), we conclude that for all $k \leq \sigma_n$,

$$\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)} = O(1/\sqrt{n}), \qquad \text{and} \qquad \hat{Y}_k^{(n)} - \hat{Y}_{k-1}^{(n)} = O(1/\sqrt{n}),$$

where in this section $o(\cdot)$ and $O(\cdot)$ are defined uniformly in space, i.e., $O(1/\sqrt{n})$ is a function which is bounded by a deterministic constant over $\sqrt{n}$ and $\sqrt{n}\, o(1/\sqrt{n})$ converges uniformly to zero as n tends to infinity.

By Theorem 7.4.1 in [14], (2.6.3) would result from the following relations,

$$\lim_{n\to\infty} \max_{1\leq k\leq\sigma_n} \left| A_k^{n,x} - h \sum_{i=0}^{k-1} \mu_x(\hat{X}_i^{(n)}, \hat{Y}_i^{(n)}) \right| = 0 \quad \text{a.s.,} \tag{2.6.4}$$

$$\lim_{n\to\infty} \max_{1\leq k\leq\sigma_n} \left| A_k^{n,y} - h \sum_{i=0}^{k-1} \mu_y(\hat{X}_i^{(n)}, \hat{Y}_i^{(n)}) \right| = 0 \quad \text{a.s.,} \tag{2.6.5}$$

$$\lim_{n\to\infty} \max_{1\leq k\leq\sigma_n} \left| A_k^{n,x,x} - \eta h \sum_{i=0}^{k-1} \sigma^2(\hat{X}_i^{(n)}, \hat{Y}_i^{(n)}) \right| = 0 \quad \text{a.s.,} \tag{2.6.6}$$

$$\lim_{n\to\infty} \max_{1\leq k\leq\sigma_n} \left| A_k^{n,y,y} - \eta(1-\rho^2) h \sum_{i=0}^{k-1} \sigma^2(\hat{X}_i^{(n)}, \hat{Y}_i^{(n)}) \right| = 0 \quad \text{a.s.,} \tag{2.6.7}$$

$$\lim_{n\to\infty} \max_{1\leq k\leq\sigma_n} |A_k^{n,x,y}| = 0 \quad \text{a.s.} \tag{2.6.8}$$

The rest of the proof is devoted to the verification of the above identities.

We start with a proof of (2.6.4). Since $\sigma^2(x, y)$ is Lipschitz continuous, (2.3.1), (2.3.2) and (2.3.11) imply that

$$|\hat{\alpha}_k - \hat{\alpha}_{k-1}| = O(\sqrt{h}).$$

In view of (2.6.1), for $k < \sigma_n$, we have,

$$-\frac{1}{2} \leq \alpha_k \leq \hat{c}(N+1),$$

for some constant $\hat{c}$. Since the event $k < \sigma_n$ is $\mathcal{F}_{k-1}$-measurable,

$$\mathbb{P}^{(n)}\left(\xi_k^X = 1 \text{ and } k < \sigma_n \mid \Xi_{k-1}\right) = \chi_{\{k<\sigma_n\}}\mathbb{P}^{(n)}\left(\xi_k^X = 1 \mid \Xi_{k-1}\right) = \chi_{\{k<\sigma_n\}}\, p_k.$$

We now use the above estimates, the definition (2.3.12) of the transition probability p_k and Taylor expansion. Then, on the set $k < \sigma_n$,

$$\begin{aligned}
&\mathbb{P}^{(n)}\left(\xi_k^X = 1 \mid \Xi_{k-1}\right)\\
&\quad= \frac{rh + \sqrt{\eta h}\ (1 + \hat{\alpha}_{k-1}\xi_{k-1}^X + \hat{\alpha}_k) - \eta h\ (1/2 + \hat{\alpha}_k) + o(h)}{2\sqrt{\eta h}\ (1 + \hat{\alpha}_k) + o(h)}\\
&\quad= \frac{rh + \sqrt{\eta h}\ (1 + \hat{\alpha}_{k-1}\xi_{k-1}^X + \hat{\alpha}_k) - \eta h\ (1/2 + \hat{\alpha}_k)}{2\sqrt{\eta h}\ (1 + \hat{\alpha}_k)} + o(\sqrt{h})\\
&\quad= \frac{1}{2} + \frac{\hat{\alpha}_{k-1}}{2(1 + \hat{\alpha}_k)}\ \xi_{k-1}^X + \frac{rh - \eta(1/2 + \hat{\alpha}_k)h}{2(1 + \hat{\alpha}_k)} + o(h).
\end{aligned} \tag{2.6.9}$$

We thus conclude that on the event $k < \sigma_n$, the following estimate holds,

$$\begin{aligned}
\mathbb{E}_{k-1}^{(n)}\left[\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)}\right] &= \sqrt{\eta h}\ \mathbb{E}_{k-1}^{(n)}\left[(1 + \hat{\alpha}_k)\xi_k^X - \hat{\alpha}_{k-1}\xi_{k-1}^X\right],\\
&= \sqrt{\eta h}\ \left[(1 + \hat{\alpha}_k)\ (2\mathbb{P}^{(n)}(\xi_k^X = 1 \mid \Xi_{k-1}) - 1) - \hat{\alpha}_{k-1}\xi_{k-1}^X\right],\\
&= rh - \eta\left(\frac{1}{2} + \hat{\alpha}_k\right)h + o(h),\\
&= \mu_x(\hat{X}_{k-1}^{(n)}, \hat{Y}_{k-1}^{(n)})h + o(h),
\end{aligned}$$

where the last equality follows from the definition of $\hat{\alpha}$, the Lipschitz continuity of $\mu(x, y)$ and (2.6.2). Then, (2.6.4) follows directly from the above estimate.

We continue with a proof of (2.6.5). We start with the definition of q_k and use the truncation introduced in (2.3.11). On $k < \sigma_n$, this fields the following estimate,

$$2 \times \mathbb{P}^{(n)}\left(\xi_k^Y = 1 \mid \Xi_{k-1}\right) - 1 = \frac{\hat{\alpha}_{k-1}}{1 + \hat{\alpha}_k}\xi_{k-1}^Y + \frac{\sqrt{h}\ \mu_y(X_{k-1}^{(n)}, Y_{k-1}^{(n)})}{\sqrt{\eta(1-\rho^2)}\ (1 + \hat{\alpha}_k)}.$$

As before we directly estimate the on $k-1 \leq \sigma_n$,

$$\begin{aligned}
&\mathbb{E}^{(n)}_{k-1}\left(\hat{Y}^{(n)}_k - \hat{Y}^{(n)}_{k-1}\right) \\
&\qquad = \sqrt{\eta(1-\rho^2)h}\ \left((1+\hat{\alpha}_k)\left(2\mathbb{P}^{(n)}\left(\xi^Y_k = 1 \mid \Xi_{k-1}\right) - 1\right) - \hat{\alpha}_k \xi^Y_{k-1}\right) \\
&\qquad = \mu_y(\hat{X}^{(n)}_{k-1}, \hat{Y}^{(n)}_{k-1})h + o(h).
\end{aligned}$$

Again, the last equality follows from (2.6.2) and the fact that $\mu_y(x,y)$ is Lipschitz continuous. This completes the proof of (2.6.5).

We continue with the quadratic estimates. Indeed, by (2.6.9), on $k < \sigma_n$,

$$2 \times \mathbb{P}^{(n)}\left(\xi^X_k = 1|\Xi_{k-1}\right) - 1 = \frac{\hat{\alpha}_{k-1}}{1+\hat{\alpha}_k}\xi^X_{k-1} + o(\sqrt{h}).$$

Since $A^{n,x}$ is predictable, on $k < \sigma_n$,

$$\mathbb{E}^{(n)}_{k-1}\left(\left(M^{n,x}_k - M^{n,x}_{k-1}\right)^2\right) = \mathbb{E}^{(n)}_{k-1}\left(\left(\hat{X}^{(n)}_k - \hat{X}^{(n)}_{k-1}\right)^2\right) + o(h)$$

and

$$\begin{aligned}
&\mathbb{E}^{(n)}_{k-1}\left(\left(\hat{X}^{(n)}_k - \hat{X}^{(n)}_{k-1}\right)^2\right) \\
&\quad = \eta h\left((1+\hat{\alpha}_k)^2 + (\hat{\alpha}_{k-1})^2 - 2\hat{\alpha}_{k-1}(1+\hat{\alpha}_k)\xi^X_{k-1}(2\mathbb{P}^{(n)}(\xi^X_k = 1|\Xi_{k-1}) - 1)\right), \\
&\quad = \eta h\ (1 + 2\hat{\alpha}^{(n)}_k), \\
&\quad = \eta h\ \sigma^2(\hat{X}^{(n)}_{k-1}, \hat{Y}^{(n)}_{k-1}),
\end{aligned}$$

and (2.6.6) follows. The relation (2.6.7) is proved similarly.

It remains to establish (2.6.8). The processes $A^{n,x}, A^{n,y}$ are predictable. Thus, from (2.3.14) it follows that, on $k < \sigma_n$,

$$\begin{aligned}
&\mathbb{E}^{(n)}_{k-1}\left(\left(M^{n,x}_k - M^{n,x}_{k-1}\right)\left(M^{n,y}_k - M^{n,y}_{k-1}\right)\right) \\
&\qquad = \mathbb{E}^{(n)}_{k-1}\left(\left(\hat{X}^{(n)}_k - \hat{X}^{(n)}_{k-1}\right)\left(\hat{Y}^{(n)}_k - \hat{Y}^{(n)}_{k-1}\right)\right) + o(h), \\
&\qquad = \mathbb{E}^{(n)}_{k-1}\left(\hat{X}^{(n)}_k - \hat{X}^{(n)}_{k-1}\right)\mathbb{E}^{(n)}_{k-1}\left(\hat{Y}^{(n)}_k - \hat{Y}^{(n)}_{k-1}\right) + o(h), \\
&\qquad = o(h),
\end{aligned}$$

where we used the fact that ξ_k^X and ξ_k^Y are conditionally independent. $\square$

2.7 Numerical Results

In this section, we present numerical results from our model for European and American Vanilla, Lookback, Geometric and Arithmetic Asian options under the Heston dynamics. Our computations are obtained by a direct implementation of the methodology described in the previous sections. In particular, we explicitly refrained from using known numerical techniques that improve the performance of the trees. This is done to ensure the replicability of our reported results.

2.7.1 Vanillas

In Tables (2.1), (2.2), and (2.3), we use the same parameter sets as in Beliaeva and Nawalkha [2], i.e. for European call and put options: strike $K = 100$; initial stock prices: $S_0 = 90, 95, 100, 105, 110$, maturities: $T = 1$ month, 3 months, and 6 months; initial volatility values: $\sqrt{\nu_0} = 0.2, 0.3, 0.4$; interest rate: $r = 0.05$; vol of vol: $\eta = 0.1$; mean reversion rate: $\kappa = 3$; long run vol: $\theta = 0.04$; and correlation: $\rho = -0.7$. For American put options: $K = 100$, $S_0 = 90, 100, 110$; $T = 1$ month, 3 months, and 6 months; $\sqrt{\nu_0} = 0.2, 0.4$; $\rho = -0.1, -0.7$; $r = 0.05$; $\eta = 0.1$; $\kappa = 3$, $\theta = 0.04$.

Tables (2.1) and (2.2) show the convergence of European put and call prices computed by our method compared to the closed form solutions of Heston [17]. In the European case, one can calculate errors as Heston's solution is available in closed form. The option prices computed for the number of time steps $N = 200$, 350 and 500 illustrate very good convergence to the closed form solutions as reported in Tables 1 and 2. Furthermore, one can verify that the put-call parity holds exactly for option prices at each of these time steps sizes. Clearly, this is the outcome of the fact that our price process in any step size is a martingale.

Table (2.3) reports the difference between the American put prices obtained from our method and those obtained by the Control Variate (CV) technique of [2]. The table shows that our numbers are in good agreement with those obtained by the CV method. The first three largest differences between the models are $(0.27\%, 0.26\%, 0.22\%)$ and on average there is a difference of 0.10% per option. We should point out to the reader that the CV technique computes the value of the put option via the formula

$$\text{CV American Price} = \text{Tree American} + (\text{Closed Form Euro - Tree Euro}).$$

According to [2], this method is particularly useful for longer maturity options.

Chockalingam and Muthuraman [6] develop a partial differential equations (PDE) based finite difference method to price American options under stochastic volatility. More specifically, they transform the free boundary problem resulting from the pricing of American options into a sequence of fixed-boundary problems of European type. The prices listed in Table (2.4) and (2.5) are taken from [6] as a benchmark for our tree based method. The authors provide the values arising from the projected successive over relaxation (PSOR) method and the component-wise splitting (CS) method. They state that other PDE based methods (see Ikonen and Toivanen [20] for a detailed analysis) fall between these two in terms of speed/accuracy and ease of implementation. As test parameters, they use the most common parameter values for American options under the Heston dynamics in the PDE-based literature: $K = 10$, $r = 0.1$, $\eta = 0.9$, $\kappa = 5.0$, $\theta = 0.16$, and $\rho = 0.1$, $T = 0.25$, $\sqrt{\nu_0} = 0.25, 0.5$. Following [6], we take the prices computed by Ikonen and Toivanen [20] (using the CS method together with a very fine grid) as the reference values. From Table (2.4) and (2.5), one can clearly conclude that our results for both $N = 250$ and $N = 350$ are very close to reference values.

2.7.2 Exotics

Our numerical experimentation confirms that backward recursion yields quite fast and accurate results for the two dimensional problems like European and American Vanilla option pricing problems. However, our numerical experimentation also reveals that the straightforward application of the recursive method takes too long on a personal computer when another continuous variable is introduced to price an exotic option. Hence, in order to substantially speed up the computations, we use our discrete equations as a discretization scheme for our Monte Carlo (MC) simulation. In other words, we carry out the MC simulation on the tree.

It is also important to note that our main concern in this section is to show the pure application of our computation method. There are many well known techniques in the literature which improve the speed and the accuracy of tree and MC methods. However, as in the backward recursion we refrain using any of these techniques.

Below we outline results for the geometric, arithmetic Asian and for lookback options.

We start with the geometric Asian and let

$$G_T = \exp\left(\frac{1}{T}\int_0^T \ln(S_t)dt\right)$$

be the geometric mean of S_t over time t during $[0, T]$. Then the payoff of a fixed strike geometric Asian call is given by $max(G_T - K, 0)$. Kim and Wee [21] provide semi-closed solutions for the price of geometric Asian options under the Heston model. We compare our results with theirs.

Table (2.6) displays a comparison between prices from the semi-closed solution and those from our MC simulation on tree with $N = 300$ and number of simulations (NumSim)= $10^5, 5 * 10^5, 10^6$. As benchmark prices, we use the values given in Table 5 from [21] for the parameter values: $S_0 = 100$, $\nu_0 = 0.09$, $r = 0.05$, $\kappa = 1.15$, $\theta = 0.348$, $\rho = -0.64$, $\eta = 0.39$. As it is clear from the table, our numerical scheme provides a very good approximation for the analytical prices. For $NumSim = 10^6$, we get three largest percentage errors as $(0.40\%, 0.34\%, 0.28\%)$ and average percentage error is 0.11%. Table (2.7) shows the 95% confidence intervals for the prices computed for different numbers of simulations.

Table (2.8) includes our results for arithmetic Asian options under the Heston model. We carry out the simulations as in the same way described previously. Let

$$A_T = \exp\left(\frac{1}{T}\int_0^T S_t dt\right)$$

be the arithmetic average of S_t over time t during $[0, T]$. Then the payoff of a fixed strike arithmetic Asian call is given by $max(A_T - K, 0)$. Pages and Printems [25] use the functional quantization based quadrature formula to price vanilla Calls and Asian Calls in the Heston model. The numbers computed from MC method, Romberg log-extrapolation, and K-interpolation of Romberg and their standard deviations in the parenthesis are tabulated for comparison (see Table 4 in [25] for a more detailed explanation of the results). We test our model using the numbers reported in their paper. As one can observe from Table (2.8), our prices together with the confidence intervals are in accordance with the only reference values for arithmetic Asian options under the Heston dynamics which can be found in the literature.

It is clear that when we price a lookback option using backward recursion, we also need another continuous variable holding the running max or min. But in this case, we can constrain this variable to take values on a tree as well. However, it still remains more efficient to apply our MC method on the tree. Table (2.9) presents numerical results obtained by the standard MC method and our numerical method for fixed strike lookback Call options. As comparison we used simple Monte Carlo simulations based on a Euler method. The table contains prices for $N = 3000$ and $NumSim = 10^5$. As one can see from the last column, the numbers obtained from our numerical method differ only slightly from the prices computed by the Euler MC method.

In terms of the theoretical complexity, we require n^3 many computations for n many time steps in the difference equations case. This is similar to that of PDE approach. More precisely, Table (2.10) provides average running times for the options in Tables 1-9. The computer used is a standard laptop with an Intel Core i7 M620@2.67 GHz CPU and a 4-GB memory. The algorithm was implemented in MATLAB.

2.8 Concluding Remarks

In this paper, we have developed a recombining tree approximation of the Heston model. Our approach is very general and applies to all stochastic volatility models with a factor equation. Low-dimensional European and the American option equation can be solved by a straightforward backward recursion. We have done extensive numerical experimentation with the resulting pricing equations. These results, reported in the previous section, confirm the efficiency of the method.

Bibliography

[1] Aldous, D. (1981). Weak convergence of stochastic processes for processes viewed in the strasbourg manner. Unpublished Manuscript, Statistics Laboratory Univ. Cambridge.

[2] Beliaeva, N.A. and Nawalkha, S.K. (2010). A simple approach to pricing American options under the Heston stochastic volatility model. *Journal of Derivatives*, 17, 25–43.

[3] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. Wiley, New York.

[4] Black, F. and Scholes, M. (1973). The valuation of options and corporate liability. *Journal of Political Economy*, 81, 637–654.

[5] Chesney, M. and Scott, L. (1989). A comparison of the modified Black-Scholes model and a random variance model. *Journal of Financial and Quantitative Analysis*, 24 267–284.

[6] Chockalingam, A. and Muthuraman, K. (2011). American Options Under Stochastic Volatility. *Operations Research*, 59(4), 793-809.

[7] Cox, J.C., Ingersoll, J.E. and Ross, S.A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53, 385–407.

[8] Cox, J.C., Ross, S.A. and Rubinstein, M. (1979). Options pricing: A simplified approach. *Journal of Financial Economics*, 7, 229–263.

[9] Dolinsky, Y. and Soner, H.M. (2012). Duality and Convergence for Binomial Markets with Friction *Finance and Stochastics*, forthcoming.

[10] Duan, Jin. (1995). The GARCH option pricing model. *Mathematical Finance* , 5, 13–132.

[11] Duan, Jin. (1996a). Cracking the smile. *Risk* , 9, 55–59.

[12] Duan, Jin. (1996b). A unified theory of option pricing under stochastic volatility-from GARCH to diffusion. Unpublished manuscript, Hong Kong University of Science and Technology.

[13] Duan, Jin. (1997). Augmented GARCH(p,q) process and its diffusion limit. *Journal of Econometrics* , 79, 97–127.

[14] Ethier, S. N. and Kurtz, T. G. (1985). *Markov Processes: Characterization and Convergence*. Wiley, New York.

[15] Gruber, U. and Schweizer, M. (2006). A Diffusion Limit for Generalized Correlated Random Walks. *Journal of Applied Probability* , 43, 60–73.

[16] Hagan, P., Kumar, D., Lesniewski, A., and Woodward, D. (2002). Managing Smile Risk. *Wilmott Mag.*, 84–108.

[17] Heston, S. (1993). A closed form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6, 329–343.

[18] Hilliard, J. E. and Schwartz, A. L. (1996). Binomial Option Pricing Under Stochastic Volatility and Correlated State Variables. *Journal of Derivatives* , 4, 23–39.

[19] Hull, J.C. and White, A. (1987). The Pricing of Options on Assets with Stochastic Volatility. *Journal of Finance*, 42, 281–300.

[20] Ikonen, S. and Toivanen, J. (2007b). Efficient numerical methods for pricing American options under stochastic volatility *Numer. Methods for Partial Differential Equations*, 24(1), 104–126.

[21] Kim, B. and Wee, I. (2011). Pricing of geometric Asian options under Heston's stochastic volatility model *Quantitative Finance*, forthcoming.

[22] Kusuoka, S. (1995). Limit Theorems on Option Replications with Transaction Costs. *Annals of Applied Probability*, 5, 198–221.

[23] Nelson, D. (1990). ARCH models as diffusion approximations. *Journal of Econometrics*,45, 7–38.

[24] Nelson, D.B. and Ramaswamy, K. (1990). Simple Binomial Process as Diffusion Approximations in Financial Models. *The Review of Financial Studies*,3, 393–430.

[25] Pages, G. and Printems, J. (2005). Functional quantization for numerics with an application to option pricing. *Monte Carlo Methods Appl.*, 11, 407–446.

[26] Peskir, G., and Shiryaev, A.N. (2006). *Optimal Stopping and Free-Boundary Problems.* Birkhauser, ETH Zurich.

[27] Ritchken, P. and Trevor, R. (1999). Pricing Options under Generalized GARCH and Stochastic Volatility Process. *Journal of Finance*, 54(1), 377–402.

[28] Rubinstein, M. (1992). "Guiding Force," in From Black-Scholes to Black Holes: New Frontiers in Options. *Risk Magazine.*

[29] Stein, E.M and Stein, J.C. (1991). Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *The Review of Financial Studies*, 4, 727–752.

Table 2.1: Convergence of European put prices versus analytical solution of Heston [17]. Parameters: $K = 100$, $r = 0.05$, $\eta = 0.1$, $\kappa = 3.0$, $\theta = 0.04$, and $\rho = -0.7$.

$S(0)$	$\sqrt{\nu}_0$	T	Tree			Analytical Solution	Error %		
			$N = 200$	$N = 350$	$N = 500$		$N = 200$	$N = 350$	$N = 500$
90	0.2	0.0833	9.6541	9.6533	9.6533	9.6533	0.01	0.00	0.00
95	0.2	0.0833	5.2059	5.2084	5.2077	5.2074	-0.03	0.02	0.01
100	0.2	0.0833	2.0953	2.0960	2.0965	2.0971	-0.08	-0.05	-0.03
105	0.2	0.0833	0.6082	0.6047	0.6050	0.6053	0.48	-0.10	-0.06
110	0.2	0.0833	0.1267	0.1271	0.1270	0.1265	0.11	0.48	0.35
90	0.3	0.0833	9.9913	9.9900	9.9900	9.9905	0.01	0.00	0.00
95	0.3	0.0833	6.0147	6.0170	6.0162	6.0155	-0.01	0.02	0.01
100	0.3	0.0833	3.1308	3.1288	3.1290	3.1302	0.02	-0.05	-0.04
105	0.3	0.0833	1.4001	1.3955	1.3955	1.3967	0.25	-0.08	-0.09
110	0.3	0.0833	0.5365	0.5374	0.5372	0.5367	-0.05	0.13	0.09
90	0.4	0.0833	10.5687	10.5670	10.5668	10.5668	0.02	0.00	0.00
95	0.4	0.0833	6.9357	6.9363	6.9352	6.9335	0.03	0.04	0.02
100	0.4	0.0833	4.1893	4.1864	4.1861	4.1852	0.10	0.03	0.02
105	0.4	0.0833	2.3280	2.3232	2.3229	2.3222	0.25	0.04	0.03
110	0.4	0.0833	1.1893	1.1897	1.1893	1.1882	0.09	0.13	0.09
90	0.2	0.25	9.5736	9.5693	9.5694	9.5698	0.04	0.00	0.00
95	0.2	0.25	5.9691	5.9685	5.9693	5.9692	0.00	-0.01	0.00
100	0.2	0.25	3.3742	3.3774	3.3794	3.3770	-0.08	0.01	0.07
105	0.2	0.25	1.7420	1.7393	1.7402	1.7410	0.06	-0.10	-0.05
110	0.2	0.25	0.8290	0.8249	0.8253	0.8259	0.37	-0.13	-0.08
90	0.3	0.25	10.5941	10.5879	10.5882	10.5893	0.04	-0.01	-0.01
95	0.3	0.25	7.3343	7.3327	7.3329	7.3316	0.04	0.02	0.02
100	0.3	0.25	4.8279	4.8331	4.8340	4.8310	-0.06	0.04	0.06
105	0.3	0.25	3.0420	3.0379	3.0391	3.0388	0.11	-0.03	0.01
110	0.3	0.25	1.8368	1.8320	1.8319	1.8325	0.23	-0.03	-0.03
90	0.4	0.25	11.8375	11.8281	11.8288	11.8287	0.07	0.00	0.00
95	0.4	0.25	8.8120	8.8081	8.8070	8.8035	0.10	0.05	0.04
100	0.4	0.25	6.3762	6.3790	6.3786	6.3735	0.04	0.09	0.08
105	0.4	0.25	4.5066	4.5005	4.5004	4.4976	0.20	0.06	0.06
110	0.4	0.25	3.1099	3.1035	3.1025	3.1011	0.28	0.08	0.05
90	0.2	0.5	9.7547	9.7545	9.7606	9.7572	-0.03	-0.03	0.04
95	0.2	0.5	6.7258	6.7248	6.7185	6.7199	0.09	0.07	-0.02
100	0.2	0.5	4.4355	4.4369	4.4320	4.4312	0.10	0.13	0.02
105	0.2	0.5	2.8077	2.8159	2.8100	2.8107	-0.11	0.18	-0.02
110	0.2	0.5	1.7286	1.7289	1.7275	1.7240	0.27	0.28	0.20
90	0.3	0.5	11.0786	11.0792	11.0845	11.0807	-0.02	-0.01	0.03
95	0.3	0.5	8.2445	8.2422	8.2367	8.2363	0.10	0.07	0.00
100	0.3	0.5	5.9835	5.9830	5.9784	5.9763	0.12	0.11	0.04
105	0.3	0.5	4.2450	4.2504	4.2449	4.2443	0.02	0.15	0.02
110	0.3	0.5	2.9647	2.9640	2.9623	2.9582	0.22	0.20	0.14
90	0.4	0.5	12.6195	12.6199	12.6231	12.6171	0.02	0.02	0.05
95	0.4	0.5	9.9373	9.9318	9.9260	9.9223	0.15	0.10	0.04
100	0.4	0.5	7.7110	7.7069	7.7017	7.6965	0.19	0.13	0.07
105	0.4	0.5	5.9065	5.9075	5.9015	5.8978	0.15	0.17	0.06
110	0.4	0.5	4.4841	4.4806	4.4779	4.4716	0.28	0.20	0.14

Table 2.2: Convergence of European call prices versus analytical solution of Heston [17]. Parameters: $K = 100$, $r = 0.05$, $\eta = 0.1$, $\kappa = 3.0$, $\theta = 0.04$, and $\rho = -0.7$.

$S(0)$	$\sqrt{\nu}_0$	T	Tree			Analytical	Error %		
			$N = 200$	$N = 350$	$N = 500$	Solution	$N = 200$	$N = 350$	$N = 500$
90	0.2	0.0833	0.0699	0.0691	0.0691	0.0691	1.13	-0.05	0.02
95	0.2	0.0833	0.6217	0.6242	0.6235	0.6232	-0.23	0.17	0.06
100	0.2	0.0833	2.5111	2.5118	2.5122	2.5129	-0.07	-0.04	-0.02
105	0.2	0.0833	6.0240	6.0205	6.0208	6.0211	0.05	-0.01	-0.01
110	0.2	0.0833	10.5425	10.5429	10.5428	10.5423	0.00	0.01	0.00
90	0.3	0.0833	0.4071	0.4058	0.4058	0.4063	0.20	-0.12	-0.12
95	0.3	0.0833	1.4305	1.4328	1.4320	1.4313	-0.06	0.10	0.05
100	0.3	0.0833	3.5466	3.5446	3.5448	3.5460	0.02	-0.04	-0.04
105	0.3	0.0833	6.8159	6.8113	6.8113	6.8125	0.05	-0.02	-0.02
110	0.3	0.0833	10.9523	10.9532	10.9530	10.9525	0.00	0.01	0.00
90	0.4	0.0833	0.9845	0.9828	0.9826	0.9826	0.19	0.02	0.00
95	0.4	0.0833	2.3515	2.3521	2.3510	2.3493	0.10	0.12	0.07
100	0.4	0.0833	4.6051	4.6022	4.6019	4.6010	0.09	0.03	0.02
105	0.4	0.0833	7.7438	7.7390	7.7387	7.7380	0.08	0.01	0.01
110	0.4	0.0833	11.6051	11.6055	11.6051	11.6040	0.01	0.01	0.01
90	0.2	0.25	0.8158	0.8115	0.8116	0.8120	0.47	-0.05	-0.05
95	0.2	0.25	2.2113	2.2107	2.2116	2.2114	-0.01	-0.03	0.01
100	0.2	0.25	4.6164	4.6196	4.6216	4.6192	-0.06	0.01	0.05
105	0.2	0.25	7.9842	7.9815	7.9824	7.9832	0.01	-0.02	-0.01
110	0.2	0.25	12.0712	12.0671	12.0675	12.0682	0.03	-0.01	-0.01
90	0.3	0.25	1.8363	1.8301	1.8305	1.8316	0.26	-0.08	-0.06
95	0.3	0.25	3.5766	3.5750	3.5751	3.5738	0.08	0.03	0.03
100	0.3	0.25	6.0701	6.0753	6.0762	6.0732	-0.05	0.03	0.05
105	0.3	0.25	9.2842	9.2802	9.2813	9.2810	0.04	-0.01	0.00
110	0.3	0.25	13.0790	13.0742	13.0741	13.0747	0.03	0.00	0.00
90	0.4	0.25	3.0797	3.0703	3.0710	3.0709	0.29	-0.02	0.00
95	0.4	0.25	5.0542	5.0503	5.0493	5.0457	0.17	0.09	0.07
100	0.4	0.25	7.6184	7.6212	7.6208	7.6157	0.04	0.07	0.07
105	0.4	0.25	10.7488	10.7428	10.7426	10.7399	0.08	0.03	0.03
110	0.4	0.25	14.3521	14.3457	14.3447	14.3433	0.06	0.02	0.01
90	0.2	0.5	2.2237	2.2235	2.2296	2.2262	-0.11	-0.12	0.15
95	0.2	0.5	4.1948	4.1938	4.1875	4.1889	0.14	0.12	-0.03
100	0.2	0.5	6.9045	6.9060	6.9010	6.9002	0.06	0.08	0.01
105	0.2	0.5	10.2767	10.2849	10.2790	10.2797	-0.03	0.05	-0.01
110	0.2	0.5	14.1976	14.1979	14.1965	14.1930	0.03	0.03	0.02
90	0.3	0.5	3.5476	3.5483	3.5535	3.5497	-0.06	-0.04	0.11
95	0.3	0.5	5.7135	5.7112	5.7057	5.7053	0.14	0.10	0.01
100	0.3	0.5	8.4525	8.4520	8.4474	8.4453	0.09	0.08	0.03
105	0.3	0.5	11.7140	11.7194	11.7140	11.7133	0.01	0.05	0.01
110	0.3	0.5	15.4337	15.4330	15.4313	15.4272	0.04	0.04	0.03
90	0.4	0.5	5.0885	5.0889	5.0921	5.0861	0.05	0.06	0.12
95	0.4	0.5	7.4063	7.4008	7.3950	7.3913	0.20	0.13	0.05
100	0.4	0.5	10.1800	10.1759	10.1707	10.1655	0.14	0.10	0.05
105	0.4	0.5	13.3755	13.3765	13.3705	13.3668	0.07	0.07	0.03
110	0.4	0.5	16.9532	16.9496	16.9469	16.9406	0.07	0.05	0.04

Table 2.3: Comparison of American put prices calculated with our method and with the control variate technique of Beliaeva and Nawalkha [2]. Parameters: $K = 100$, $r = 0.05$, $\eta = 0.1$, $\kappa = 3.0$, $\theta = 0.04$, and $\rho = -0.7$.

$S(0)$	ρ	$\sqrt{\nu_0}$	T	Tree $N = 250$	Control Variate $N = 200$	Difference %
90	-0.1	0.2	0.0833	10.0000	10.0000	0.00
100	-0.1	0.2	0.0833	2.1236	2.1254	-0.08
110	-0.1	0.2	0.0833	0.1090	0.1091	-0.05
90	-0.7	0.2	0.0833	10.0000	9.9997	0.00
100	-0.7	0.2	0.0833	2.1249	2.1267	-0.08
110	-0.7	0.2	0.0833	0.1273	0.1274	-0.07
90	-0.1	0.4	0.0833	10.7123	10.7100	0.02
100	-0.1	0.4	0.0833	4.2194	4.2158	0.08
110	-0.1	0.4	0.0833	1.1666	1.1667	-0.01
90	-0.7	0.4	0.0833	10.6843	10.6804	0.04
100	-0.7	0.4	0.0833	4.2183	4.2140	0.10
110	-0.7	0.4	0.0833	1.1942	1.1939	0.02
90	-0.1	0.2	0.25	10.1713	10.1706	0.01
100	-0.1	0.2	0.25	3.4729	3.4747	-0.05
110	-0.1	0.2	0.25	0.7726	0.7736	-0.13
90	-0.7	0.2	0.25	10.1222	10.1206	0.02
100	-0.7	0.2	0.25	3.4790	3.4807	-0.05
110	-0.7	0.2	0.25	0.8405	0.8416	-0.13
90	-0.1	0.4	0.25	12.1880	12.1819	0.05
100	-0.1	0.4	0.25	6.5023	6.4964	0.09
110	-0.1	0.4	0.25	3.0952	3.0914	0.12
90	-0.7	0.4	0.25	12.1245	12.1122	0.10
100	-0.7	0.4	0.25	6.4989	6.4899	0.14
110	-0.7	0.4	0.25	3.1512	3.1456	0.18
90	-0.1	0.2	0.5	10.6521	10.6478	0.04
100	-0.1	0.2	0.5	4.6531	4.6473	0.12
110	-0.1	0.2	0.5	1.6857	1.6832	0.15
90	-0.7	0.2	0.5	10.5682	10.5637	0.04
100	-0.7	0.2	0.5	4.6691	4.6636	0.12
110	-0.7	0.2	0.5	1.7899	1.7874	0.14
90	-0.1	0.4	0.5	13.3279	13.3142	0.10
100	-0.1	0.4	0.5	8.0231	8.0083	0.18
110	-0.1	0.4	0.5	4.5554	4.5454	0.22
90	-0.7	0.4	0.5	13.2431	13.2172	0.20
100	-0.7	0.4	0.5	8.0204	7.9998	0.26
110	-0.7	0.4	0.5	4.6328	4.6201	0.27

Table 2.4: American put prices determined with our tree approach and finite difference methods. Parameters: $K = 10$, $r = 0.1$, $\eta = 0.9$, $\kappa = 5.0$, $\theta = 0.16$, and $\rho = 0.1$, $T = 0.25$, $\sqrt{\nu_0} = \mathbf{0.25}$

Method	Grid Size	S_0				
		8	9	10	11	12
PSOR	(40,16,8)	2.0000	1.0952	0.4966	0.2042	0.0838
	(60,32,66)	2.0000	1.1037	0.5142	0.2105	0.0815
	(120,64,130)	2.0000	1.1064	0.5182	0.2126	0.0819
	(240,128,258)	2.0000	1.1071	0.5193	0.2133	0.0820
Componentwise splitting	(40,16,8)	2.0004	1.1003	0.4991	0.2035	0.0828
	(60,32,66)	2.0000	1.1043	0.5147	0.2104	0.0813
	(120,64,130)	2.0000	1.1066	0.5183	0.2126	0.0819
	(240,128,258)	2.0000	1.1073	0.5194	0.2133	0.0820
Transformation procedure	(40,16,8)	2.0000	1.0952	0.4966	0.2042	0.0838
	(60,32,66)	2.0000	1.1035	0.5142	0.2105	0.0815
	(120,64,130)	2.0000	1.1063	0.5181	0.2126	0.0819
	(240,128,258)	2.0000	1.1071	0.5193	0.2133	0.0820
Our Tree Method	N					
	150	2.0000	1.1086	0.5155	0.2140	0.0825
	250	2.0000	1.1079	0.5190	0.2140	0.0822
	350	2.0000	1.1074	0.5193	0.2134	0.0828
Reference Value		2.0000	1.1076	0.5200	0.2137	0.0820

Table 2.5: American put prices determined with our tree approach and finite difference methods. Parameters: $K = 10$, $r = 0.1$, $\eta = 0.9$, $\kappa = 5.0$, $\theta = 0.16$, and $\rho = 0.1$, $T = 0.25$, $\sqrt{\nu_0} = \mathbf{0.5}$

Method	Grid Size	S_0				
		8	9	10	11	12
PSOR	(40,16,8)	2.0691	1.3139	0.7720	0.4293	0.2324
	(60,32,66)	2.0760	1.3292	0.7908	0.4442	0.2405
	(120,64,130)	2.0775	1.3320	0.7940	0.4467	0.2419
	(240,128,258)	2.0779	1.3329	0.7951	0.4476	0.2424
Componentwise splitting	(40,16,8)	2.0676	1.3094	0.7646	0.4232	0.2297
	(60,32,66)	2.0758	1.3287	0.7900	0.4435	0.2401
	(120,64,130)	2.0774	1.3317	0.7936	0.4463	0.2417
	(240,128,258)	2.0780	1.3328	0.7949	0.4474	0.2423
Transformation procedure	(40,16,8)	2.0691	1.3140	0.7721	0.4294	0.2325
	(60,32,66)	2.0760	1.3291	0.7908	0.4442	0.2405
	(120,64,130)	2.0775	1.3319	0.7940	0.4467	0.2419
	(240,128,258)	2.0780	1.3329	0.7951	0.4476	0.2424
Our Tree Method	N					
	150	2.0791	1.3362	0.7957	0.4495	0.2435
	250	2.0786	1.3338	0.7964	0.4501	0.2435
	350	2.0790	1.3339	0.7964	0.4485	0.2440
Reference Value		2.0784	1.3336	0.7960	0.4483	0.2428

Table 2.6: Comparison of our method and the semi-closed solution for fixed-strike geometric Asian call options for : $S_0 = 100$, $\nu_0 = 0.09$, $r = 0.05$, $\kappa = 1.15$, $\theta = 0.348$, $\rho = -0.64$, $\eta = 0.39$.

T	K	MC on Tree with $N = 300$			Semi-Closed Solution	Difference %		
		NumSim				NumSim		
		10^5	$5 * 10^5$	10^6		10^5	$5 * 10^5$	10^6
0.2	90	10.6598	10.6551	10.6562	10.6571	0.02	-0.02	-0.01
0.2	95	6.6006	6.5970	6.5888	6.5871	0.20	0.15	0.03
0.2	100	3.4699	3.4564	3.4510	3.4478	0.64	0.25	0.09
0.2	105	1.4697	1.4610	1.4611	1.4552	1.00	0.40	0.40
0.2	110	0.4730	0.4742	0.4719	0.4724	0.14	0.38	-0.10
0.4	90	11.7310	11.7111	11.7077	11.7112	0.17	0.00	-0.03
0.4	95	8.0988	8.1067	8.0877	8.0894	0.12	0.21	-0.02
0.4	100	5.1480	5.1746	5.1641	5.1616	-0.26	0.25	0.05
0.4	105	3.0414	3.0060	3.0040	3.0018	1.32	0.14	0.07
0.4	110	1.5555	1.5776	1.5679	1.5715	-1.02	0.39	-0.23
0.5	90	12.2974	12.2495	12.2330	12.2329	0.53	0.14	0.00
0.5	95	8.7711	8.7668	8.7753	8.7553	0.18	0.13	0.23
0.5	100	5.9036	5.9151	5.9008	5.8971	0.11	0.31	0.06
0.5	105	3.7150	3.7120	3.7165	3.7072	0.21	0.13	0.25
0.5	110	2.1622	2.1692	2.1595	2.1589	0.15	0.48	0.03
1	90	14.5646	14.6087	14.5937	14.5779	-0.09	0.21	0.11
1	95	11.6287	11.5518	11.5474	11.5551	0.64	-0.03	-0.07
1	100	8.9708	8.9378	8.9530	8.9457	0.28	-0.09	0.08
1	105	6.8003	6.7392	6.7505	6.7559	0.66	-0.25	-0.08
1	110	5.0161	4.9878	4.9704	4.9722	0.88	0.31	-0.04
1.5	90	16.3889	16.4588	16.5200	16.5030	-0.69	-0.27	0.10
1.5	95	13.7324	13.7764	13.7690	13.7625	-0.22	0.10	0.05
1.5	100	11.3599	11.3247	11.3304	11.3374	0.20	-0.11	-0.06
1.5	105	9.2487	9.2187	9.2076	9.2245	0.26	-0.06	-0.18
1.5	110	7.4342	7.3959	7.4019	7.4122	0.30	-0.22	-0.14
2	90	18.0757	18.1112	18.0816	18.0914	-0.09	0.11	-0.05
2	95	15.6133	15.6021	15.5211	15.5640	0.32	0.24	-0.28
2	100	13.3624	13.3245	13.2833	13.2933	0.52	0.24	-0.08
2	105	11.2855	11.2862	11.2627	11.2728	0.11	0.12	-0.09
2	110	9.4243	9.4840	9.4901	9.4921	-0.71	-0.09	-0.02
3	90	20.6523	20.4276	20.5149	20.5102	0.69	-0.40	0.02
3	95	18.3985	18.2361	18.2884	18.3060	0.51	-0.38	-0.10
3	100	16.2151	16.2555	16.2609	16.2895	-0.46	-0.21	-0.18
3	105	14.5000	14.4330	14.4046	14.4531	0.32	-0.14	-0.34
3	110	12.6065	12.8177	12.7982	12.7882	-1.42	0.23	0.08

Table 2.7: Confidence Intervals for fixed-strike geometric Asian call options for : $S_0 = 100$, $\nu_0 = 0.09$, $r = 0.05$, $\kappa = 1.15$, $\theta = 0.348$, $\rho = -0.64$, $\eta = 0.39$.

Confidence Intervals 95 %		
$NumSim = 10^5$	$NumSim = 5 * 10^5$	$NumSim = 10^6$
$(10.6135, 10.7060)$	$(10.6345, 10.6758)$	$(10.6416, 10.6708)$
$(6.5609, 6.6402)$	$(6.5793, 6.6147)$	$(6.5763, 6.6014)$
$(3.4397, 3.5001)$	$(3.4429, 3.4699)$	$(3.4415, 3.4605)$
$(1.4501, 1.4894)$	$(1.4522, 1.4698)$	$(1.4548, 1.4673)$
$(0.4623, 0.4838)$	$(0.4694, 0.4790)$	$(0.4685, 0.4753)$
$(11.6678, 11.7941)$	$(11.6829, 11.7394)$	$(11.6877, 11.7277)$
$(8.0438, 8.1538)$	$(8.0820, 8.1313)$	$(8.0703, 8.1051)$
$(5.1027, 5.1932)$	$(5.1543, 5.1948)$	$(5.1498, 5.1784)$
$(3.0065, 3.0764)$	$(2.9904, 3.0216)$	$(2.9930, 3.0150)$
$(1.5308, 1.5803)$	$(1.5665, 1.5887)$	$(1.5601, 1.5758)$
$(12.2270, 12.3679)$	$(12.2181, 12.2808)$	$(12.2108, 12.2552)$
$(8.7094, 8.8328)$	$(8.7391, 8.7944)$	$(8.7557, 8.7949)$
$(5.8516, 5.9556)$	$(5.8919, 5.9384)$	$(5.8843, 5.9172)$
$(3.6735, 3.7566)$	$(3.6934, 3.7306)$	$(3.7034, 3.7297)$
$(2.1305, 2.1938)$	$(2.1551, 2.1833)$	$(2.1495, 2.1694)$
$(14.4642, 14.6650)$	$(14.5638, 14.6536)$	$(14.5619, 14.6255)$
$(11.5367, 11.7208)$	$(11.5109, 11.5927)$	$(11.5186, 11.5763)$
$(8.8888, 9.0528)$	$(8.9013, 8.9744)$	$(8.9272, 8.9789)$
$(6.7282, 6.8724)$	$(6.7072, 6.7713)$	$(6.7278, 6.7732)$
$(4.9538, 5.0784)$	$(4.9601, 5.0154)$	$(4.9508, 4.9899)$
$(16.2635, 16.5144)$	$(16.4023, 16.5152)$	$(16.4800, 16.5599)$
$(13.6150, 13.8498)$	$(13.7239, 13.8289)$	$(13.7319, 13.8061)$
$(11.2523, 11.4676)$	$(11.2765, 11.3729)$	$(11.2963, 11.3645)$
$(9.1503, 9.3471)$	$(9.1749, 9.2626)$	$(9.1766, 9.2387)$
$(7.3457, 7.5226)$	$(7.3563, 7.4355)$	$(7.3739, 7.4299)$
$(17.9261, 18.2253)$	$(18.0442, 18.1782)$	$(18.0342, 18.1289)$
$(15.4721, 15.7544)$	$(15.5392, 15.6651)$	$(15.4767, 15.5654)$
$(13.2303, 13.4945)$	$(13.2656, 13.3835)$	$(13.2416, 13.3249)$
$(11.1626, 11.4084)$	$(11.2315, 11.3409)$	$(11.2240, 11.3013)$
$(9.3113, 9.5373)$	$(9.4334, 9.5345)$	$(9.4544, 9.5258)$
$(20.4610, 20.8436)$	$(20.3429, 20.5123)$	$(20.4547, 20.5750)$
$(18.2156, 18.5814)$	$(18.1549, 18.3174)$	$(18.2310, 18.3459)$
$(16.0417, 16.3885)$	$(16.1781, 16.3329)$	$(16.2062, 16.3157)$
$(14.3343, 14.6656)$	$(14.3594, 14.5067)$	$(14.3525, 14.4566)$
$(12.4514, 12.7617)$	$(12.7476, 12.8878)$	$(12.7489, 12.8476)$

Table 2.8: Comparison of our method and the Functional Quantization Method by Pages and Printems [25] for Arithmetic Asian options. Parameters: $S_0 = 50$, $\nu_0 = 0.01$, $r = 0.05$, $\kappa = 2$, $\theta = 0.01$, $\rho = 0.5$, $\eta = 0.1$.

K	10^8 - MC	Crude MC Reference	Romberg on crude FQ		K-interpol of Romberg		Our Method $N = 300, NumSim = 10^6$ Price	Our Method $N = 300, NumSim = 10^6$ Conf. Int.
44	6.92	(0.08%)	6.92	(0.01%)	6.92	(0.01%)	6.9196	(6.9139 , 6.9252)
45	5.97	(0.10%)	5.97	(0.04%)	5.97	(0.02%)	5.9768	(5.9712 , 5.9825)
46	5.03	(0.11%)	5.03	(0.05%)	5.03	(0.02%)	5.0334	(5.0278 , 5.0390)
47	4.11	(0.14%)	4.12	(0.09%)	4.11	(0.04%)	4.1117	(4.1062 , 4.1172)
48	3.245	(0.16%)	3.25	(0.17%)	3.24	(0.05%)	3.2506	(3.2453 , 3.2559)
49	2.46	(0.20%)	2.47	(0.32%)	2.46	(0.04%)	2.4673	(2.4624 , 2.4723)
50	1.79	(0.26%)	1.80	(0.63%)	1.79	(0.03%)	1.7926	(1.7882 , 1.7970)
51	1.25	(0.31%)	1.26	(1.16%)	1.25	(0.17%)	1.2541	(1.2503 , 1.2580)
52	0.84	(0.39%)	0.85	(2.06%)	0.84	(0.37%)	0.8430	(0.8398 , 0.8463)
53	0.54	(0.50%)	0.56	(3.73%)	0.545	(0.78%)	0.5502	(0.5475 , 0.5529)
54	0.34	(0.63%)	0.36	(6.58%)	0.34	(1.37%)	0.3485	(0.3464 , 0.3506)
55	0.21	(0.81%)	0.23	(11.53%)	0.21	(2.15%)	0.2159	(0.2142 , 0.2176)
56	0.125	(1.04%)	0.15	(19.96%)	0.125	(2.84%)	0.1317	(0.1303 , 0.1330)

Table 2.9: Comparison of our method and Euler Simulation for Lookback Call option with fixed strike. Parameters: $S_0 = 100$, $\nu_0 = 0.16$, $r = 0.05$, $\kappa = 3$, $\theta = 0.04$, $\rho = -0.7$, $\eta = 0.1$.

T	K	Euler Simulation $n = 3000$, $NumSim = 10^5$		Our method $N = 3000$, $NumSim = 10^5$		Difference %
		price	Confidence Interval	price	Confidence Interval	
0.2	90	23.4527	(23.3844 , 23.5210)	23.4679	(23.3996 , 23.5362)	0.06
0.2	95	18.5511	(18.4827 , 18.6196)	18.5459	(18.4776 , 18.6142)	0.03
0.2	100	13.5145	(13.4464 , 13.5825)	13.6562	(13.5878 , 13.7246)	1.05
0.2	105	9.2629	(9.1987 , 9.3272)	9.2620	(9.1978 , 9.3262)	0.01
0.2	110	6.0746	(6.0185 , 6.1306)	6.0899	(6.0340 , 6.1457)	0.25
0.4	90	27.7252	(27.6333 , 27.8172)	27.7378	(27.6461 , 27.8296)	0.05
0.4	95	22.7931	(22.7015 , 22.8846)	22.7784	(22.6869 , 22.8698)	0.06
0.4	100	17.8937	(17.8017 , 17.9857)	17.9052	(17.8136 , 17.9969)	0.06
0.4	105	13.5301	(13.4415 , 13.6187)	13.6541	(13.5649 , 13.7434)	0.92
0.4	110	10.0038	(9.9224 , 10.0852)	10.0978	(10.0160 , 10.1796)	0.94
0.5	90	29.1737	(29.0738 , 29.2735)	29.2407	(29.1405 , 29.3409)	0.23
0.5	95	24.2728	(24.1733 , 24.3722)	24.3094	(24.2095 , 24.4093)	0.15
0.5	100	19.4547	(19.3542 , 19.5552)	19.5036	(19.4033 , 19.6038)	0.25
0.5	105	15.1074	(15.0099 , 15.2049)	15.0772	(14.9801 , 15.1742)	0.20
0.5	110	11.4637	(11.3730 , 11.5544)	11.4401	(11.3498 , 11.5305)	0.21
1	90	34.1211	(33.9910 , 34.2511)	34.1944	(34.0646 , 34.3242)	0.21
1	95	29.4579	(29.3273 , 29.5886)	29.4015	(29.2720 , 29.5311)	0.19
1	100	24.6878	(24.5573 , 24.8184)	24.7163	(24.5855 , 24.8470)	0.12
1	105	20.1960	(20.0686 , 20.3234)	20.3721	(20.2443 , 20.4999)	0.87
1	110	16.5429	(16.4206 , 16.6652)	16.4579	(16.3367 , 16.5791)	0.51
1.5	90	37.6113	(37.4587 , 37.7640)	37.8563	(37.7035 , 38.0091)	0.65
1.5	95	33.2861	(33.1314 , 33.4408)	33.0959	(32.9428 , 33.2491)	0.57
1.5	100	28.5915	(28.4380 , 28.7451)	28.3913	(28.2386 , 28.5440)	0.70
1.5	105	24.2427	(24.0913 , 24.3941)	24.1616	(24.0107 , 24.3124)	0.33
1.5	110	20.4593	(20.3131 , 20.6054)	20.4385	(20.2919 , 20.5850)	0.10
2	90	41.0722	(40.8963 , 41.2481)	41.0605	(40.8861 , 41.2350)	0.03
2	95	36.6204	(36.4454 , 36.7953)	36.5932	(36.4185 , 36.7680)	0.07
2	100	31.9362	(31.7612 , 32.1112)	32.0618	(31.8874 , 32.2361)	0.39
2	105	27.8954	(27.7220 , 28.0688)	27.7302	(27.5578 , 27.9026)	0.59
2	110	24.0406	(23.8719 , 24.2093)	23.8907	(23.7223 , 24.0591)	0.62
3	90	47.0043	(46.7881 , 47.2205)	47.0854	(46.8698 , 47.3010)	0.17
3	95	42.6606	(42.4453 , 42.8759)	42.5750	(42.3599 , 42.7901)	0.20
3	100	38.6746	(38.4588 , 38.8903)	38.3630	(38.1469 , 38.5790)	0.81
3	105	34.5038	(34.2898 , 34.7177)	34.2793	(34.0657 , 34.4929)	0.65
3	110	30.7339	(30.5229 , 30.9449)	30.4407	(30.2312 , 30.6502)	0.95

Table 2.10: Average Running Times for options in Tables 1-9

		time in seconds
European Put and Call (Table 1 & 2)	$N = 200$ $N = 350$ $N = 500$	5.71 30.37 89.27
American Put (Table 3)	$N = 250$	13.97
American Put (Table 4 & 5)	$N = 150$ $N = 250$ $N = 350$	3.15 14.66 40.50
Geometric Asian (Table 6 & 7)	$NumSim = 10^5$ $NumSim = 5 * 10^5$ $NumSim = 10^6$	8.17 40.91 81.79
Arithmetic Asian (Table 8)	$N = 300$ $NumSim = 10^6$	98.65
Lookback (Table 9)	$N = 3000$ $NumSim = 10^5$	94.54

Chapter 3

Dual Currency Credit Default Swap: Theoretical and Empirical Analysis

Erdinç Akyıldırım, Loriano Mancini, and Emrah Şener

3.1 Introduction

As markets remain uncertain over the debt dynamic sustainability of eurozone peripheral countries, sovereign CDSs are being widely adopted as an investment and risk management tool. Divergent investor demand for these contracts in USD vs EUR format has led to the evolution of an active quanto market in sovereign CDS. This paper is an empirical study of the properties of violations of the law of one price in the European CDS market around period of market distress. We investigate the extent to which limits to arbitrage are state-dependent and how they have been affected during crises.

A question, which is not attacked widely in literature, is whether CDSs across local and foreign currencies should be identical. Under the assumption of zero correlation between default rates and exchange rates[1], the equality of CDS rates across two currencies can in fact be proven. This, however, is empirically not the case. During the recent 2010-2013 period, we observed a major anomaly in European CDS markets, where spreads between EUR and USD CDSs have witnessed record magnitudes. In April 2010, the price of EUR CDS, referencing Germany, was nearly 29% cheaper than the same contract denominated

[1] Throughout the paper, USDEUR will denote EUR per unit of USD.

in USD, given that this value was only 7% in February 2010[2]. During the same period, even more dramatic spreads were observed for Greece. Then, far from reflecting a view on the default of Germany - a sovereign with almost no default risk - the fact that this country's spreads expanded considerably gives an idea on how negative sentiments across the continent (or even across globe) are creeping into the credit-related products that settle in EUR currency. During an unstable period, concerns over European defaults, such as those over Greece, may have caused a major fall in the demand for EUR-denominated protection. The investor knew that if say, Greece were to default, then her 1-million EUR worth of protection would be worth much less at the time of default, due to a potential major devaluation of EUR. In these circumstances, the investor would understandably be less willing to get into such a contract. But still, price spreads being surprisingly large and persistent raise serious doubts regarding the validity of the law of one price. This motivates the following questions: 1) What type of behavioural and/or fundamental drivers appeared to have caused the corresponding CDS prices to drift apart so significantly for long periods of time? 2) Should the USD CDS really be used as a *de facto* variable for pricing the same credit risk denominated in EUR? 3) Under the law of one price, is it possible to interpret the dynamics of currency dependence by a mere correlation factor?

The law of one price (LOP) requires that two assets generating equal cash-flows have equal values. Otherwise, an arbitrage opportunity arises. Arbitrageurs are essential in eliminating mispricings, providing liquidity to economy, and allowing asset prices to reflect market fundamentals. While it is not common to observe significant deviations from the law of one price, the recent US credit crisis and European turbulence offer a valuable opportunity to analyse this important phenomena. While it is not very surprising to gather scattered evidence of violations of the LOP in small and illiquid markets, it is perhaps more intriguing to investigate the properties of such violations in large and liquid markets. In doing so, we study six European markets that quote CDS both in USD and EUR: Belgium, France, Germany, Ireland, Italy, and Spain.

Our investigation is three-fold. First, we calculate a proxy of deviation from the LOP called *Quanto CDS*. This is based on the spreads between two CDS prices issued by the same sovereign but denominated in different currencies. In doing so, we take the correlation

[2] From Fitch Solutions (2010)

between default risk and exchange rate into consideration and form a variable that should be zero if the LOP holds. Then, considering the whole sample data (August 2010 to June 2013) which includes the Europe specific turbulence periods, we investigate whether our proxy steered and persistently fluctuated away from zero. Second, we conduct a factor analysis which is useful in reducing the dimension by concentrating on a few important factors that represent the main sources of variation in the dual currency CDS market. We find that the first two factors statically significant during the entire period of the time. Thirdly, in light of these results, we suggest and test a model in pricing dual currency CDSs in European markets, which in effect helps us to understand currency dependence. This is an important question since many financial institutions are tempted to employ USD as the *de facto* currency to price EUR denominated products of the same issuer. The motivation is that USD spread is a perfect substitute to discount risks denominated in other currencies given that risk premia (of the same issuer) across two currencies is identical, which obviously is not satisfied in segmented markets.

Our empirical results are as follows. First, we observe that our proxy of deviation from the LOP is considerably large for each sovereign, suggesting that correlation is not sufficient in explaining currency dependence in European CDS markets. We also see that such deviations are state-dependent. More specifically, during calmer periods, the fluctuations of LOP deviations are comparably milder than distressful periods. Nonetheless, Spain displays the highest LOP divergence *after* 2010, which coincides with the country's credit rating down grades by the rating agencies. Similar observations can also be made for currency dependence dynamics. Our second finding is related to the price discovery in the corresponding CDSs, which sheds light to the natural habitat of traders. It must be noted that dollar-denominated CDSs always tend to lead the price variations of EUR-denominated CDSs for all countries. Our findings also suggest that the more developed a European country is, the stronger its USD CDS contribution on EUR CDS. This may be suggestive of the fact that the natural habitat of arbitrageurs may be different depending on the currency denomination used to fund operations in different markets. Our third finding is concerned more on whether USD spread can in fact be used as a *de facto* input for the pricing of the same credit risk denominated in EUR. Although this seems to be a convenient approach, investors require different premia on USD and EUR denominated

CDSs of the same issuer.

In light of these arguments, we empirically test a correlation based diffusion model with a jump component. We observe that even if the issuer is the same, it is empirically incorrect to treat USD spread as the reference currency to price credit products denominated in other currencies. Accurate pricing models should consider the different risk dynamics of the corresponding currencies and include in one way or another currency-dependent factors. Obviously, in segmented markets, risk management and pricing should require an understanding of how individual risks and assets interact instead of only focusing on a single risk factor.

The paper is organized as follows: Section (3.2) outlines the results in the literature that are related to our work. Section (3.3) gives the framework for our theoretical and empirical investigation of the dynamics of the LOP deviation proxy. Section (3.4) details the data selection. Section (3.5) discusses the factor analysis. Section (3.6) provides price discovery in the corresponding CDSs. Section (3.7) includes the stochastic model and (3.8) provides the econometric estimation of this model and finally Section (3.9) concludes.

3.2 Literature Review

Our paper is related to credit risk for which a vast theoretical body of work exists. Merton [47], Black and Cox [19], and Leland [43] present structural models in pricing and analyzing defaultable corporate bonds. On the other hand, Jarrow and Turnbull [39] , Artzner and Delbaen [18] , Duffie and Singleton [27], and Elliott et al. [30] take a different route and discuss what is called reduced-form models. In addition, to address the idea that a firm can default instantaneously after a sudden drop in its value, Zhou [50] incorporates jump risks into corporate bond models and argues that a jump component makes it highly flexible to generate various shapes of risk premium term structure, explaining certain empirical regularities, such as recovery rates and default probabilities. Similarly, Cariboni and Schoutens [24] discuss how CDSs can be priced under a more general Levy model [3], where default occurs at the time a pre-set barrier is crossed. In practice, it is common to see investment banks, data suppliers (i.e. Reuters and Bloomberg), and rating

[3] Refer to Sato [49] for a detailed account of Levy processes.

agency companies (i.e. Moody's and Standard Poors) to use USD as a *de facto* variable in pricing credit products referencing the same issuer, even across different currencies (Merrill Lynch, 2000). The main motivation arises from the liquidity and the depth of the exchange rate market. In a *single* currency setting, Elton et al. [31] decompose corporate credit spreads into three components: default, tax and risk premia. Collin-Dufresne et al. [26] discusses the existence of an unexplained systemic factor affecting credit spreads. Hull et al. [37] derives a theoretical relationship among credit risk premiums and credit default swaps, proving that they must be equal under no-arbitrage condition. Longstaff et al. [44] decompose credit spreads into default and non-default components, giving empirical evidence for the significant impact of default components. Other works such as Duffee [29], Driessen [28], Ericsson and Renault [33] , Chen et al. [25], and Feldhatter and Lando [35] also divide corporate credit spreads into various components. Buraschi et al. [23] argue that credit spreads may be linked to the systemic component as a consequence of heterogeneous perceptions of market agents. In the *multiple* currencies setting, Domowitz et al. [9] discuss the relationship of market volatility and country/currency risk premiums. Kercheval et al. [41] derive an arbitrage condition, where the changes in credit risk premiums across two currencies are correlated. Their findings give evidence for the existence of currency dependence in the cross-currency bonds of various cross-market entities (i.e. Toyota, Dresdner Bank, and European Investment Bank). They argue that the changes in the risk premiums may be attributed to several changes other than the perceived creditworthiness of the issuer. Ehlers and Schönbucher [32] argue that the differentials in credit risk premiums are driven by the correlation between default risk and exchange rate, as well as by the jumps in the exchange rate market. Finally, Jankowitsch and Pichler [38] demonstrate why credit risk premiums should be equal across different currencies when default and exchange rates are independent from each other. Their findings also indicate an existence of currency dependence in international corporate bonds. Landschoot [42] argues how corporate credit spreads in two different currencies of two *different* issuers respond to various market factors.

Second, this paper contributes to the large body of empirical literature on financial market integration. In fully integrated markets, only risk factors to be priced are global risks (e.g., Solnik [15], Adler and Dumas [16]) while in segmented markets only local risks are priced. (e.g., Stulz [14], Bekaert and Harvey [13]). Several recent papers provide evidence that sovereign credit markets and equity markets and are becoming more integrated within the world market (e.g., De Jong and De Roon [12], Pukthuanthong and Roll [11], and Longstaff et al. [45]). This suggests that credit markets are increasingly driven by global rather than by local factors. Further, several international asset pricing models (e.g.,Stulz [14], Dumas and Solnik [10]) show that currency risk factors could play an important role for asset returns and they show that, in equilibrium, investors require to be compensated for bearing exchange rate risk. De Santis and Gerard [8] , Vassalou [7], and Lustig and Verdelhan [3] report empirical evidence of a premium for currency risk.

3.3 Theoretical Framework

In order to explore the concept of *Quanto CDS* in the sovereign eurozone credit markets, we consider the credit default swaps of the the same issuer. The only difference between these two markets is the currency denomination of the CDS (USD and EUR, respectively). First, similar to Buraschi et al. [22] and Kercheval et al. [41], we consider a simple two-period setting and derive a no-arbitrage relationship, decomposing it into two parts: (a) the CIRP component; (b) the credit spread component.

According to the covered interest rate parity (CIRP), the following condition must hold for a riskless investment between period t and T:

$$(1 + R^d(t,T)) = \frac{X(t)}{F(t,T)}(1 + R^e(t,T)), \tag{3.3.1}$$

where $R^i(t,T)$ is the arithmetic risk-free rate in the two corresponding currencies $i = (d,e)$, being USD and EUR, respectively, and $X(t)$ and $F(t,T)$ are the EUR/USD (EUR per USD) spot and forward exchange rates, respectively. The idea is that an investor who borrows 1 dollar today, thus owing $(1 + R^d(t,T))$ at time T, can convert 1 dollar to $X(t)$ euros at time t, invest them in EUR deposits, thus receiving EUR $X(t)(1 + R^e(t,T))$ at maturity. If the forward exchange rate is $F(t,T)$, the dollar value today of this investment is $X(t)(1 + R^e(t,T))/F(t,T)$, which needs therefore to equate $(1 + R^d(t,T))$ unless an arbitrage opportunity exists.

One can extend the previous argument to sovereign defaultable bonds. Consider that

Spain, for instance, issues two pure discount bonds with maturity T in two different currencies (i.e. USD and EUR). If the bond can only default at the time the face value is due, i.e. at time T, then if δ^i and $S^i(t,T)$ are the recovery rates and arithmetic credit spreads in the corresponding currencies, then the following condition must also hold in a frictionless market:

$$\delta^d \times (1 + R^d(t,T) + S^d(t,T)) = \frac{X(t)}{F(t,T)}(1 + R^e(t,T) + S^e(t,T)) \times \delta^e. \tag{3.3.2}$$

The same argument of the CIRP applies.[4] If the recovery rates across foreign bonds is the same, i.e. $\delta^d = \delta^e$, by simple algebra we can rearrange Eq. (3.3.2) as follows:

$$(1 + R^d(t,T)) + S^d(t,T) = \frac{X(t)}{F(t,T)}(1 + R^e(t,T)) + \frac{X(t)}{F(t,T)}S^e(t,T). \tag{3.3.3}$$

If (3.3.1) holds, then

$$\underbrace{(1 + R^d(t,T)) - \frac{X(t)}{F(t,T)}(1 + R^e(t,T))}_{By\ CIRP\ =\ 0} + \left[S^d(t,T) - \frac{X(t)}{F(t,T)}S^e(t,T)\right] = 0. \tag{3.3.4}$$

The first component represents the CIRP condition and the second component is related to the specific pricing of the two credit spreads. This naturally implies that if CIRP holds, a necessary condition for no-arbitrage is that $S^d(t,T) - \frac{X(t)}{F(t,T)}S^e(t,T) = 0$, which leads to the following simple condition on credit spreads:

$$\frac{S^e(t,T)}{S^d(t,T)} = \frac{1 + R^e(t,T)}{1 + R^d(t,T)} \tag{3.3.5}$$

This implies that by no-arbitrage the quoted ratio of credit spreads in different currencies must be equal to the ratio of their respective risk-free rates.[5] The intuition is simple. The face value of the bond denominated in the highest interest rates currency is subject to a higher expected depreciation. This expected loss needs to be compensated ex-ante by a larger spread. The higher the *euro* risk-free rate, the higher the *euro* yield spread.

[4] One can borrow 1 dollar of an issuer's USD-denominated bond, exchange it to X_t EUR, buy X_t euros of the same issuer's EUR-denominated bond, and enter a forward contract to convert $X_t(1 + R^e(t,T) + S^e(t,T))$ euros to dollars at F(t,T).

[5] According to this relationship, if the risk-free rates are unequal but time-variant, the spreads will thus be unequal but still perfectly correlated.

Assuming CIRP violations do not exactly offset the spread no-arbitrage violations, a sufficient condition for credit mispricing is therefore:

$$S^e(t,T) - \frac{F(t,T)}{X(t)} S^d(t,T) \neq 0 \tag{3.3.6}$$

3.3.1 Empirical Implementation

Given the previous results and under the traditional textbook assumption of no institutional frictions of (3.3.2), we can investigate the deviation from the Law of One Price by studying the expected profits:

$$QuantoCDS = \underbrace{\frac{F_{bid}(t,T)}{X_{ask}(t)} S^d_{bid}(t,T) - S^e_{ask}(t,T)}_{CreditSpread\ Mispricing} \tag{3.3.7}$$

We explicitly consider the bid/ask quotes to incorporate the transactions costs. In absence of frictions, these conditions are no-arbitrage conditions, where the profits are expected to be zero.

Note also that T is not restricted to any specific maturity and hence can take on any finite value across the term structure. One of the most important implications of (3.3.4) is that it must hold for any T. Therefore, one can easily deduce that there exists at least one arbitrage opportunity if the equality is violated for at least one T. If the equality is violated for *every* T, we have a continuous term structure of covered interest LOP deviations. If the LOP holds, then the term-structure of LOP violations is naturally flat at the zero axis. If not, however, a dynamic term-structure appears at every t, though not necessarily with a monotonically positive or negative slope. One of the main objectives of this paper is indeed to investigate the nature of this term-structure at every point t in order to study whether the dynamics of violations vary significantly across different maturity points.

3.4 Data Description

We obtain data on CDS spreads from Markit Group Ltd. which follows a strict procedure to collect daily quotes of CDS spreads from major dealers in the credit derivatives market and creates composite CDS spreads. Our data set covers the period 20 August 2010 to 12 June 2013 for three-, five-, seven-, ten-year USD and EUR denominated CDS spreads for six eurozone countries which are Belgium, France, Germany, Italy, Ireland, and Spain. The differentials between USD and EUR CDS spreads before August 2010 are very close zero. Hence we conduct our empirical investigation starting from this date.

As a visual inspection, the behaviour of the five-year Quanto CDS spreads for each sovereign is plotted in the figures below. However, the figures are similar for the other maturities as well. It can be observed that Quanto CDSs are time-varying and state-dependent. It becomes highly volatile during distressful periods. Starting from the end of 2010 and until the mid of 2012, there is an increasing trend in the spreads and the volatility of deviations increase for all the sovereigns. Starting from the mid of 2012, deviations smooth down and markets start to converge back to theoretical expectations under LOP. As it is clear from the figures, the Quanto CDS spreads are higher for economically distressed counties like Ireland, Italy, and Spain.

Table (3.1) and Table (3.2) show the summary statistics for the term structure of CDS spreads denominated both in USD and EUR and Quanto CDSs for each country. It can be seen that currency adjusted differentials of five year CDS spreads for each sovereign are significantly large and the differentials are on average the largest for Spain, and the lowest for Germany. The same observation holds across different maturities. The Quanto CDSs are closely related to the expected default intensities of two sovereigns, where the risk is relatively higher for Spain. The maximum Quanto differentials are also informative about how severely the credit risk of the same issuer may change under different currencies.

The term structure of the data set also allows us to make further analysis about the economic outlook of the countries. For example, as time to maturity increases, the average of both USD and EUR CDS spreads also increase for economically stronger countries like Germany, France, and Belgium as expected. This observation also holds for the Quanto

CDS of these countries. However, for economically distressed countries like Spain and Ireland the average CDS spreads decrease across maturity and average Quanto CDS levels remain flat. Another interesting observation is that although Ireland has CDS levels higher than Spain, Spain has the highest Quanto CDS which may explain why market participants consider Spain as the systematically most important country together with Italy.

Table 3.1: Summary Statistics
This table shows the means, minimums, and maximums of the three-, five-, seven-, ten-year USD and EUR denominated CDS spreads and also Quanto CDS spreads for Belgium, France, Germany. The data set covers the period 20 August 2010 to 12 June 2013

		Country								
		Belgium			**France**			**Germany**		
	Maturity	**Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**
USD CDS	**3 Years**	143.1	19.4	392.5	85.6	16.1	214.2	34.7	6.2	91.8
	5 Years	173.7	56.0	404.4	122.0	59.3	247.3	59.4	24.1	115.7
	7 Years	182.9	84.9	401.6	137.1	71.0	255.6	72.7	36.6	125.0
	10 Years	187.9	100.4	397.5	148.6	77.5	263.8	83.5	42.3	134.7
EUR CDS	**3 Years**	105.3	9.2	330.9	55.7	7.8	169.8	19.9	2.9	65.1
	5 Years	125.2	31.8	339.3	77.9	30.8	172.4	34.3	9.5	80.5
	7 Years	131.4	49.0	335.1	88.5	44.2	205.8	42.9	18.1	88.0
	10 Years	134.2	55.8	329.5	96.8	53.0	214.0	50.4	24.6	95.2
Quanto CDS	**3 Years**	37.8	2.9	78.8	30.0	-2.0	80.8	14.8	3.1	39.7
	5 Years	48.5	17.3	90.8	44.0	8.8	99.6	25.1	8.1	66.6
	7 Years	51.5	3.0	93.5	48.6	0.1	106.1	29.8	9.3	81.2
	10 Years	53.7	2.6	95.6	51.8	0.4	108.0	33.1	9.2	88.0

Table 3.2: Summary Statistics
This table shows the means, minimums, and maximums of the three-, five-, seven-, ten-year USD and EUR denominated CDS spreads and also Quanto CDS spreads for Ireland, Italy, and Spain. The data set covers the period 20 August 2010 to 12 June 2013

		Country								
		Ireland			**Italy**			**Spain**		
	Maturity	**Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**
USD CDS	**3 Years**	560.3	87.4	1427.7	287.7	95.2	613.8	315.0	143.8	626.6
	5 Years	521.1	136.2	1263.4	315.2	124.8	590.6	338.9	196.6	633.5
	7 Years	500.7	164.9	1202.0	320.4	131.3	576.4	342.3	201.4	616.9
	10 Years	469.6	178.9	1127.3	320.0	137.8	558.8	339.7	204.6	592.1
EUR CDS	**3 Years**	507.0	65.0	1356.5	235.8	59.3	530.6	246.7	112.1	503.3
	5 Years	462.9	101.0	1194.0	256.3	78.7	501.5	264.5	135.7	504.2
	7 Years	441.5	126.2	1130.1	259.4	84.4	486.3	266.2	141.1	485.6
	10 Years	410.0	135.7	1050.5	257.6	90.0	467.9	262.4	142.5	460.4
Quanto CDS	**3 Years**	53.3	3.3	102.9	52.0	22.6	92.1	68.3	27.7	128.0
	5 Years	58.2	21.8	106.0	58.8	32.5	94.2	74.4	35.8	135.3
	7 Years	59.3	7.8	112.3	61.1	36.9	95.9	76.2	39.1	137.2
	10 Years	59.7	10.6	117.1	62.4	38.8	100.9	77.4	42.8	137.5

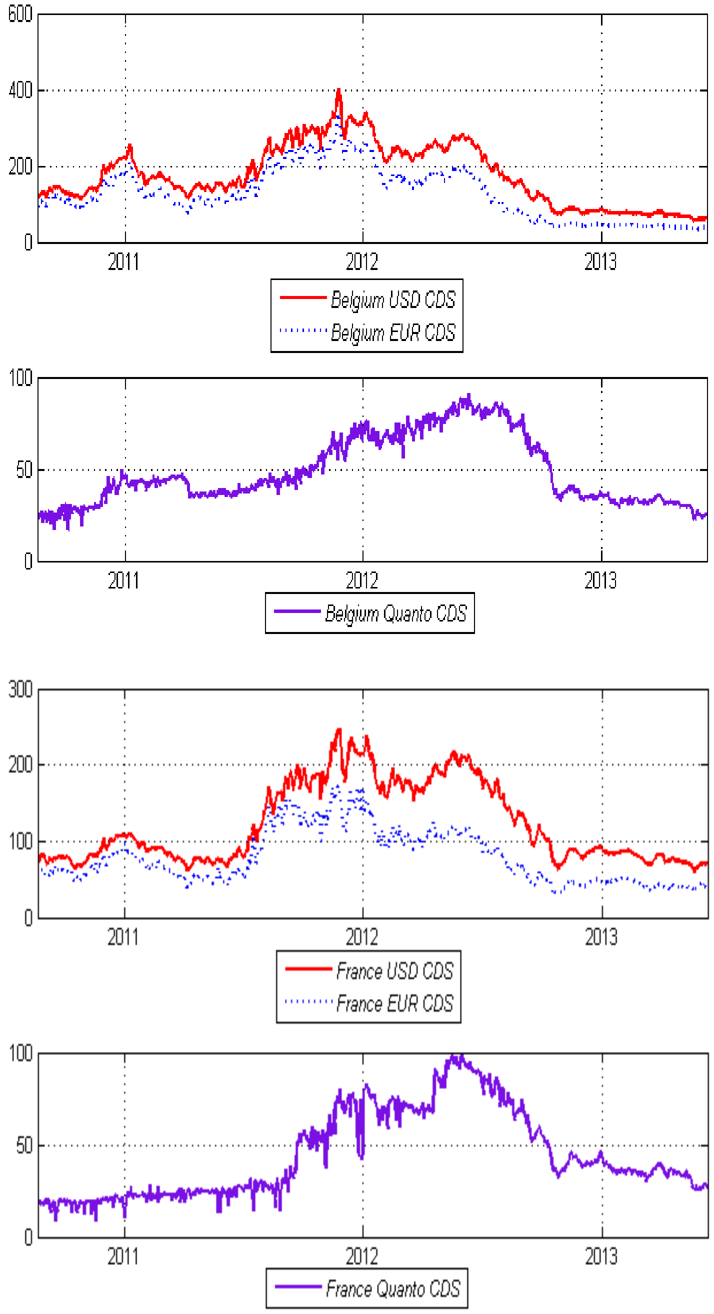

Figure 3.1: *USD, EUR and Quanto CDS spreads*
The figures above display the evolution of five-year USD, EUR and Quanto CDS spreads for the countries Belgium and France for the period 20 August 2010 to 12 June 2013.

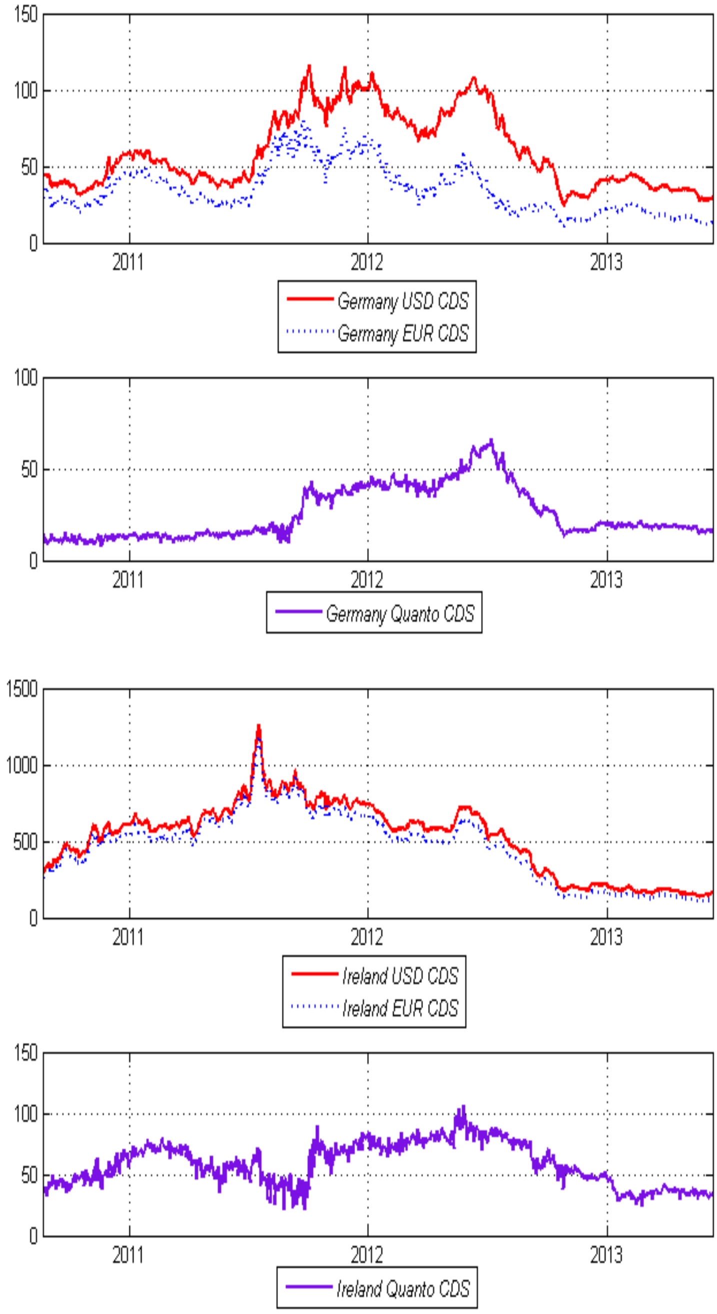

Figure 3.2: *USD, EUR and Quanto CDS spreads*
The figures above display the evolution of five-year USD, EUR and Quanto CDS spreads for the countries Germany and Ireland for the period 20 August 2010 to 12 June 2013.

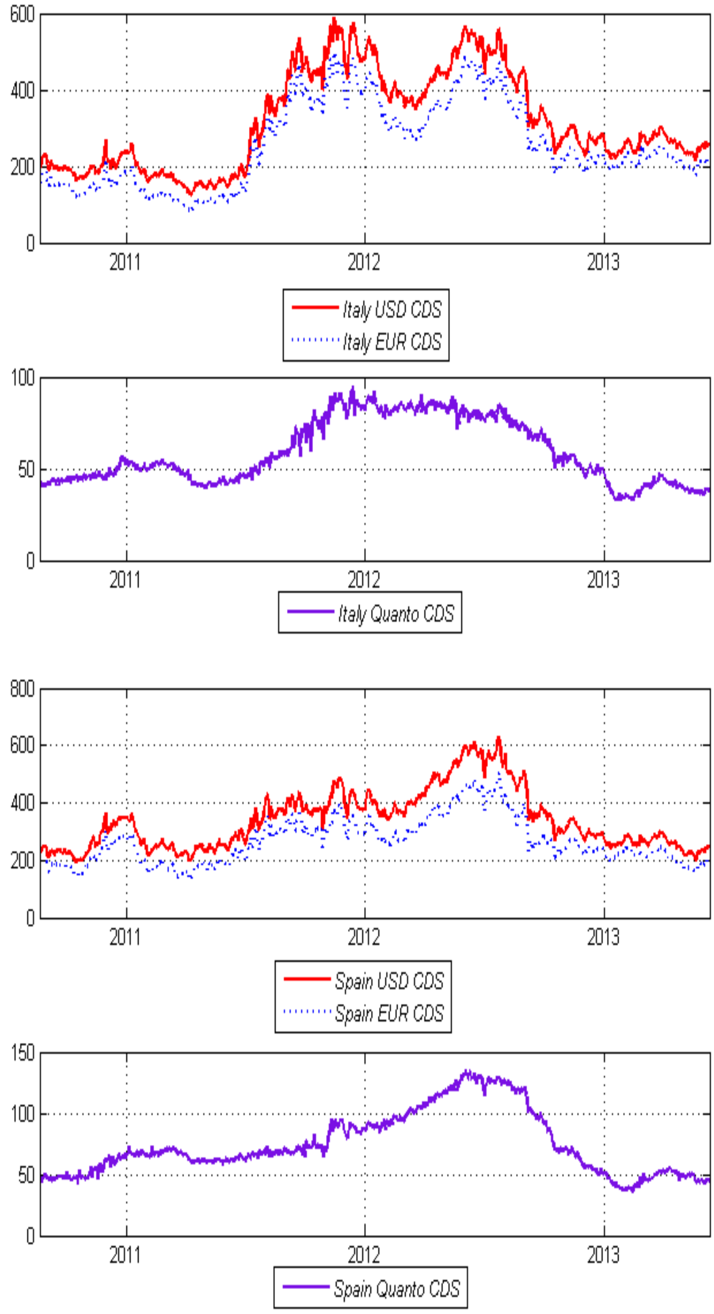

Figure 3.3: *USD, EUR and Quanto CDS spreads*
The figures above display the evolution of five-year USD, EUR and Quanto CDS spreads for the countries Italy and Spain for the period 20 August 2010 to 12 June 2013.

3.5 Commonality in Dollar and Euro Credit Risk Premiums

Are Dollar and Euro credit risk premiums affected by the same risk factors across different economies? If they are, to what extent does this co-movement occur? In order to give an answer to these questions, we conduct factor analysis.

Factor analysis is useful in "reducing the dimension" by concentrating on a few important factors that represent the main sources of variation in the market. The first component often refers to a systemic factor and the second component refers to some local (idiosyncratic) factors. In the spirit of Avellaneda and Scherer [6] we conduct both static and dynamic factor analysis. Static analysis corresponds to calculating the components for the whole sample. In the dynamic version, principal components are performed on sub-samples and also consecutive windows. The factor analysis is applied on all sub-periods separately, allowing for the measure of the evolution of important market shifts. For robustness check, we observe the variations of the coupling component (percent of variance explained by the first component) across time and we form a dynamic picture of the behaviour of the European markets.[6] In the spirit of Lustig et al. [1], we call these first two factors as the global and local factors, respectively. In other words, the eigenvector with the largest eigenvalue corresponds to the variance attributable to global risk and the second component strongly suggests the existence of a volatility risk factor associated to a country specific factor.

Table (3.3) presents the static factor results conducted separately on Dollar CDS, Euro CDS, and Quanto CDS. The table is further divided with respect to different maturities. In our unreported results we find the first two factors statically significant during the entire period of the time[7]. First we observe the coupling component, namely a global factor across time to form a dynamic picture of how the USD and EUR credit risk premia behave.[8] Following Avellaneda and Scherer [6] we propose the following categories: *Extreme*

[6]For the dynamic factor analysis, we use rolling window periods of 60 days.

[7]In order to decide the number of eigenvectors on which we should attach significance, we run a sphericity test as proposed by Fluery (1988, Ch. 2). Our sphericity test statistics show that there are two factors that are stable and interpretable

[8]Scherer and Avellaneda (2002) [6] define the coupling coefficient as "the fraction of variance attributed

Coupling, *Strong Coupling* and *Weak Coupling*.[9]

Table (3.3) reveals that there is a strong coupling in for the global credit risk component. As for the USD CDS and EUR CDS, we observe that the coupling is always *strong*: (around 79%) and *strong* (around 74%), respectively for the Quanto CDS the coupling is always *extreme*: (around 83%). This suggests that for credit risk premia denominated in different currencies, almost the entire impact arises from global systematic risks. This result validates the findings of Pan and Singleton [48]. Moreover, Longstaff et al. [45] show that the CDS spreads around the globe share a strong common relation to global financial variables. This result is also important since it demonstrates how common dependence of this type could induce significant correlations across sovereign credit spreads. Our dynamic factor results also show that the impact of a systemic shock is consistently and mostly pronounced during the market turmoil.

---

to the first component". More specifically, if $\lambda_1 > \lambda_2 > \ldots > \lambda_n$ are the eigenvalues, then

$$\text{Coupling Coefficient} = \frac{\lambda_1}{\sum_{i=1}^{n} \lambda_i}. \tag{3.5.1}$$

The coupling coefficient is a statistical representation of how often the markets co-move as a single block.

[9] *Extreme Coupling* (percentage of variance explained by the first principal component is above 80%), *Strong Coupling*, (percentage of variance explained by the first principal component is between 65-80%), and finally *Weak Coupling* (percentage of variance explained by the first principal component is less than 65%).

Table 3.3: Factor Analysis on CDS spreads
This table presents the static factor results conducted separately on Dollar CDS, Euro CDS and Quanto CDS for three-, five-, seven-, ten-year maturities.

		Maturity							
	Principal	*3 Year*		*5 Year*		*7 Year*		*10 Year*	
	Components	**Exp.**	**Cum.**	**Exp.**	**Cum.**	**Exp.**	**Cum.**	**Exp.**	**Cum.**
USD CDS	**Factor 1**	79.3%	79.3%	79.9%	79.9%	78.5%	78.5%	77.5%	77.5%
	Factor 2	13.1%	92.4%	14.8%	94.7%	16.2%	94.7%	16.9%	94.4%
	Factor 3	5.3%	97.7%	3.5%	98.2%	3.3%	98.0%	3.3%	97.7%
EUR CDS	**Factor 1**	74.6%	74.6%	73.9%	73.9%	72.2%	72.2%	70.9%	70.9%
	Factor 2	16.1%	90.7%	18.9%	92.8%	19.9%	92.1%	20.6%	91.5%
	Factor 3	5.8%	96.5%	4.7%	97.5%	4.8%	96.9%	4.9%	96.4%
Quanto CDS	**Factor 1**	83.6%	83.6%	87.6%	87.6%	84.9%	84.9%	82.5%	82.5%
	Factor 2	8.7%	92.3%	6.9%	94.5%	8.1%	93.0%	9.7%	92.1%
	Factor 3	3.3%	95.6%	2.4%	96.9%	3.1%	96.2%	3.8%	95.9%

3.6 Price Discovery

The sovereign credit market is often driven by differences in demand for CDS in the two currencies. The natural bid for USD protection on European sovereigns is often from non-European investors, total return investors or investors taking a macro view. However, many eurozone-based investors hedging sovereign exposure prefer to have their assets and liabilities in the same currency and thus provide the natural bid for EUR CDS protection. Since the order flow of the CDS markets are fragmented, the price discovery may be different depending either on the institutional characteristics of the markets in which most trade occur, or the currency habitat of the traders, or the funding currency of the trading operations. If risk premiums are not identical and reflect risk factors that are not entirely a function of the characteristics of the issuer, some important questions emerge. Does a CDS in one currency provide more timely information than its equivalent in the other currency? Which CDS retains a greater contribution to price discovery? Being able to answer these questions is important since they are related to which credit spread should be used in modelling sovereign default risk.

As defined by Lehmann (2002), price discovery can be described as timely incorporation of the trading activities into market prices. To this end, we set out to investigate the differences in the information contents of credit risks denominated in different currencies. There are two traditional ways to conduct price discovery analysis. The first one is based on the information share (IS) measure, as suggested by Hasbrouck [36]. The second one is based on the component share (CS) measure, as suggested by Gonzalo and Granger [5].[10] Following Blanco et al. [20], we calculate the IS measures to find the contribution of USD credit risk premiums to EUR credit risk premiums. We also follow Eun and SabherWal [4] in generating two spread series by using the midpoint of the last bid and ask quotes in each market.[11] We first test for the existence of cointegration across credit spreads of a single issuer and estimate a VEC (Vector Error Correction) model.[12] Given our objective, we conduct the analysis for the entire period without subdividing the sample. Then, we

[10] Both measures rely on the estimation of a vector error-correction models (VECM) of market prices; but IS assumes that price volatility reflects new information, and it allows for the correlation among multiple markets via the variance and covariance of price innovations.

[11] They prefer using quotes over transaction prices, since transaction prices may suffer from the autocorrelation problem arising due to infrequent trading.

[12] We used Johansen's cointegration test, proposed by Johansen [40], and determined the number of lags according to the Akaike information criterion. We find that the USD and EUR credit risk premiums are co-integrated whole sample period.

estimate the relative contribution to price discovery of each of the two bond spreads by computing Hasbrouck's IS price discovery measures. Since the LOP in frictionless markets implies that $S^d - c_1 \times S^e = 0$, with $c_1 \equiv (1 + R^d)/(1 + R^e)$, we estimate the following VECM specification:

$$\Delta S_t^d = A_1(S_{t-1}^d - c_1 S_{t-1}^e) + \sum_{n=1}^{N} \phi_{1n} \Delta S_{t-n}^d + \sum_{n=1}^{N} \gamma_{1n} \Delta S_{t-n}^e + u_{1t} \tag{3.6.1}$$

$$\Delta S_t^e = A_2(S_{t-1}^d - c_1 S_{t-1}^e) + \sum_{n=1}^{N} \phi_{2n} \Delta S_{t-n}^d + \sum_{n=1}^{N} \gamma_{2n} \Delta S_{t-n}^e + u_{2t} \tag{3.6.2}$$

S^d and S^e are the USD and EUR credit risk premiums, respectively, and u_{1t} and u_{2t} are the error terms. In our specification, the term $(S_{t-1}^d - c_1 S_{t-1}^e)$ captures the no-arbitrage relation that needs to be satisfied in a frictionless economy. In presence of frictions, if the EUR risk premium contributes to price discovery, then A_1 should be statistically significant. On the other hand, if the USD risk premium contributes to price discovery, then A_2 should be statistically significant. If both coefficients are significant, then both credit markets are important in the price discovery process.

Table (3.4) summarizes the results for the upper, lower and averages of the Hasbrouck bounds. If the average of the Hasbrouck bounds is greater (less) than 0.5 then USD (EUR) credit risk premium leads EUR (USD) credit risk premium. Our findings reveal that for Germany, France, Italy and Spain CDS the A_2 coefficient in ((3.6.2)) is positive and statistically significant, implying that the USD-denominated CDS strictly lead their EUR-denominated equivalents in terms of price discovery.

Most data suppliers (i.e. Reuters and Bloomberg) and rating agencies (i.e. Moody's and Standard Poors) use USD-denominated CDSs as inputs in modelling the EUR-denominated risk premium curves of the same issuer. The market convention is to use USD as de-facto currency in the valuation of foreign denominated credit products. Our results suggest that this approach might be incorrect. Not only investors seem to require different compensations for the variations in USD- and EUR-denominated CDSs but also the price discovery process differ across different CDS markets. This seems related to their specific funding markets. The assumption of currency-independence for credit spread of the same issuer may lead to seriously mispriced credit products. Validating the findings of Breger [2], risk models for credit products denominated in different currencies than USD may require assumptions that are specific to the currency of denomination.[13]

Table 3.4: Price Discovery Common Factor Results
This table summarizes the results for the upper, lower and averages of the Hasbrouck bounds. If the average of the Hasbrouck bounds is greater (less) than 0.5 then USD (EUR) credit risk premium leads EUR (USD) credit risk premium.

	A1	**A2**	**S(x)**	**S(y)**	**SM**
France	-0.03	0.13	0.38	0.98	0.68
	[-0.86306]	[3.30216]			
Germany	0.01	0.19	0.41	1.00	0.70
	[0.10714]	[3.58859]			
Italy	0.27	0.65	0.15	0.98	0.56
	[1.24914]	[3.14663]			
Spain	-0.04	0.1	0.11	0.98	0.54
	[-0.68494]	[1.67704]			
Belgium	0.02	0.15	0.37	0.99	0.68
	[0.25500]	[1.74375]			
Ireland	0.02	0.36	0.98	0.99	0.98
	[0.41261]	[4.29644]			

* The parameters named A1 and A2 in table are the coefficients in the Vector Error Correction Model, given that A1 corresponds to USD CDS and A2 corresponds to EUR CDS. The t-statistics are shown immediately below. The relevant results can be viewed under *SM* column, which is the mean of two Hasbrouck measures.

3.7 Stochastic Model

In our factor analysis in Section (3.5), we find that two factors represent the main sources of variation in USD and EUR CDS spread differentials. Hence, in the light of our previous findings, in this section, we employ a stochastic model which considers the correlation between default intensity and exchange rate and the jump in the exchange rate at the time of default as two important factors in CDS pricing. As discussed in Jankowitsch and Pichler [38], the correlation between default intensity and exchange rate is important in explaining the differentials between CDSs under two currencies, issued by the same

[13] Breger [2] decomposes the bond excess returns of the same issuer (denominated in different currencies) into four elements, and highlights that credit risk models for these bonds must be built independently. Their model is as follows

$$r_{excess} = (r_{IR} + r_{curr} + r_{spread} + r_{specific})$$

where r_{IR} denotes the changes in interest rates, r_{curr} denotes the changes in currency exchange rates, r_{spread} denotes the changes in credit spreads and $r_{specific}$ denotes the specific factors not explained by common factors.

sovereign. However, one remaining question is whether the correlation is sufficient to understand these differentials.

We use the model developed by Ehlers and Schönbucher [32] and Ehlers [17] and follow their procedure closely. However, our method differs from their method at several points. First of all, in Section (3.6) we find that for most of the sovereigns under consideration USD-denominated CDSs strictly lead their EUR-denominated equivalents in terms of price discovery. Hence, we obtain empirical evidence for taking USD as our numeraire currency in the valuation of foreign denominated credit products. Moreover, Longstaff et al. [44] provide the closed form solutions for CDS prices (by applying the standard results from Duffie et al. [27]) when the default intensity follows a square root process. Hence, in order to increase the accuracy and speed of the computations, we also employ the analytical formulas for CDS pricing derived by Longstaff et al. [44]. Furthermore, in our numerical experimentation we observe that approximating CDS spreads with the formula $(1 - R)\lambda_t$ does not yield highly accurate estimation of the model parameters. As a remedy, we use a numerical optimization technique to invert for the default intensities from the 5-year CDS spreads given the other parameter values. We also provide parameter estimates under both domestic and foreign spot martingale measures together with market prices of risk parameters thanks to our robust maximum likelihood estimation application. Finally, we apply the model with the five-year CDS spread data for six European sovereigns which have totally different characteristics than Japanese corporate obligors.[14]

We assume that the correlations between the exchange rate and interest rates are zero, and the correlations between the default intensities and interest rates are zero. We also decompose the exchange rate dynamics into three components: idiosyncratic diffusion, default intensity diffusion and a one-time jump process. Let $X = \{X_t\}_{t\geq 0}$ represent the exchange rate process (domestic currency per unit of foreign currency) under the Domestic Spot Martingale Measure (DSMM), $\mathbb{Q}_d := \mathbb{Q}_\$$, and let $r_d := r_\$$ and $r_f := r_€$ be the local and foreign interest rates, respectively. The default intensity $\lambda^{Q_\$}$ under $\mathbb{Q}_\$$ is assumed to follow an affine process, which avoids negative values. Hence we have the following model

[14] Ehlers [17] investigates the Japanese firms in the empirical section.

under $\mathbb{Q}_\$$,

$$\begin{aligned} \frac{dX}{X_-} &= (r_t^\$ - r_t^€)dt + \gamma_1\sqrt{\lambda^{Q_\$}}dW^1 + \gamma_2 dW^2 - \delta(dN - \lambda^{Q_\$}dt), && (3.7.1)\\ d\lambda^{Q_\$} &= \kappa^{Q_\$}(\theta^{Q_\$} - \lambda^{Q_\$})dt + \sigma^{Q_\$}\sqrt{\lambda^{Q_\$}}dW^1, && (3.7.2) \end{aligned}$$

with $\gamma_2 \geq 0$, $\gamma_1 \in R$ and $\delta < 1$ and W^1, W^2 are standard independent BMs under $\mathbb{Q}_\$$ and $M_t := N_t - \lambda^{Q_\$}t$ is a compensated martingale under $\mathbb{Q}_\$$. This model assumes that the foreign currency depreciates against domestic currency by a fraction of $(1-\delta)$ upon default i.e. $X_\tau = (1-\delta)X_{\tau^-}$. It can be seen that the correlation is integrated into the framework by allowing the noise and square root of default intensity to drive the exchange rate dynamics. Since the data is generated under the physical measure $\mathbb{P}$, for tractability it is assumed that the state price density is of the form

$$\left.\frac{d\mathbb{P}}{d\mathbb{Q}_\$}\right|_{\mathcal{F}_T} := L_T \quad \text{with} \quad \frac{dL}{L_-} = \phi_1\sqrt{\lambda^{Q_\$}}dW^1 + \phi_2 dW^2 - \Phi(dN - \lambda^{Q_\$}dt), \qquad (3.7.3)$$

for $\phi_1, \phi_2 \in \mathbb{R}$ and $\Phi < 1$. Under this change of measure, the processes defined by

$$\begin{aligned} \widetilde{W}^1(t) &:= W^1(t) - \phi_1\int_0^t \sqrt{\lambda^{Q_\$}(s)}ds, && (3.7.4)\\ \widetilde{W}^2(t) &:= W^2(t) - \phi_2 t, && (3.7.5) \end{aligned}$$

are standard $\mathbb{P}$-BMs and $\widetilde{M}(t) := M(t) + \Phi\lambda^\$ t$ is a compensated martingale under $\mathbb{P}$. Then the dynamics of the default intensity process $\lambda^{Q_\$}$ in terms of $\widetilde{W}^1(t)$ can be written as

$$\begin{aligned} d\lambda^{Q_\$}(t) &= \kappa^{Q_\$}(\theta^{Q_\$} - \lambda^{Q_\$}(t))dt + \sigma^{Q_\$}\sqrt{\lambda^{Q_\$}(t)}\ d\left(\widetilde{W}^1(t) + \phi_1\int_0^t \sqrt{\lambda^{Q_\$}(s)}ds\right),\\ &= (\kappa^{Q_\$} - \phi_1\sigma^{Q_\$})\left(\frac{\kappa^{Q_\$}\theta^{Q_\$}}{\kappa^{Q_\$} - \phi_1\sigma^{Q_\$}} - \lambda^{Q_\$}(t)\right)dt + \sigma^{Q_\$}\sqrt{\lambda^{Q_\$}(t)}d\widetilde{W}^1(t). \end{aligned}$$

From Proposition 11 in Ehlers [17], $\widetilde{\lambda}^{Q_\$}$ (the default intensity under the physical measure) and $\lambda^{Q_\$}$ (default intensity under DSMM, $\mathbb{Q}_\$$) satisfy the relationship

$$\widetilde{\lambda}^{Q_\$} = (1-\Phi)\lambda^{Q_\$},$$

if a default were to happen at time t. Then the dynamics of $\widetilde{\lambda}^{Q_\$}$ in case of a jump is

$$\begin{aligned} d\widetilde{\lambda}^{Q_\$}(t) &= (1-\Phi)(\kappa^{Q_\$}-\phi_1\sigma^{Q_\$})\left(\frac{\kappa^{Q_\$}\theta^{Q_\$}}{\kappa^{Q_\$}-\phi_1\sigma^{Q_\$}}-\lambda^{Q_\$}(t)\right)dt \\ &\quad +(1-\Phi)\sigma^{Q_\$}\sqrt{\lambda^{Q_\$}(t)}d\widetilde{W}^1(t), \\ &= (\kappa^{Q_\$}-\phi_1\sigma^{Q_\$})\left(\frac{(1-\Phi)\kappa^{Q_\$}\theta^{Q_\$}}{\kappa^{Q_\$}-\phi_1\sigma^{Q_\$}}-\widetilde{\lambda}^{Q_\$}(t)\right)dt \\ &\quad +\sqrt{(1-\Phi)}\sigma^{Q_\$}\sqrt{\widetilde{\lambda}^{Q_\$}(t)}d\widetilde{W}^1(t). \end{aligned}$$

Let $\kappa^p, \theta^p, \sigma^p$ denote the parameters under the physical measure $\mathbb{P}$. Then

$$d\widetilde{\lambda}^{Q_\$} = \kappa^p(\theta^p - \widetilde{\lambda}^{Q_\$}(t))dt + \sigma^p\sqrt{\widetilde{\lambda}^{Q_\$}(t)}d\widetilde{W}^1(t), \tag{3.7.6}$$

together with

$$\kappa^p := (\kappa^{Q_\$}-\phi_1\sigma^{Q_\$}), \tag{3.7.7}$$

$$\theta^p := \frac{(1-\Phi)\kappa^{Q_\$}\theta^{Q_\$}}{\kappa^{Q_\$}-\phi_1\sigma^{Q_\$}}, \tag{3.7.8}$$

$$\sigma^p := \sqrt{(1-\Phi)}\sigma^{Q_\$}. \tag{3.7.9}$$

Now if we write the exchange rate process (3.7.1) in terms of $\widetilde{W}^1(t)$, $\widetilde{W}^2(t)$, and $\widetilde{M}(t)$,

$$\begin{aligned} \frac{dX}{X_-} &= (r_t^{\$}-r_t^{€})dt + \gamma_1\sqrt{\lambda^{Q_\$}}d\widetilde{W}^1(t) + \gamma_1\phi_1\lambda^{Q_\$}dt \\ &\quad +\gamma_2 d\widetilde{W}^2(t) + \gamma_2\phi_2 dt - \delta(d\widetilde{M}(t) - \Phi\lambda^{Q_\$}dt). \end{aligned} \tag{3.7.10}$$

As the dual CDS modelling and pricing will be based on both domestic and foreign currencies for a particular sovereign, it is necessary to work under both domestic and foreign martingale measures. For this purpose, we define Foreign Spot Martingale Measure

(FSMM), $\mathbb{Q}_{€}$, by using the following change of measure

$$\frac{d\mathbb{Q}_{€}}{d\mathbb{Q}_{\$}}\Big|_{\mathcal{F}_T} := L_T \quad with \quad \frac{dL}{L_-} = \gamma_1\sqrt{\lambda^{Q_\$}}dW^1 + \gamma_2 dW^2 - \delta(dN - \lambda^{Q_\$}dt). \tag{3.7.11}$$

Then the processes defined by

$$\begin{aligned}\widehat{W}^1(t) &:= W^1(t) - \gamma_1\int_0^t \sqrt{\lambda^{Q_\$}(s)}ds,\\ \widehat{W}^2(t) &:= W^2(t) - \gamma_2 t,\end{aligned}$$

are standard $\mathbb{Q}_{€}$ BMs.

The dynamics of the default intensity process $\lambda^{Q_\$}$ in terms of $\widehat{W}^1(t)$ can be written as

$$\begin{aligned}d\lambda^{Q_\$}(t) &= \kappa^{Q_\$}(\theta^{Q_\$} - \lambda^{\$}(t))dt + \sigma^{\$}\sqrt{\lambda^{Q_\$}(t)}d\left[\widehat{W}^1(t) + \gamma_1\int_0^t\sqrt{\lambda^{Q_\$}(s)}ds\right],\\ &= (\kappa^{Q_\$} - \gamma_1\sigma^{Q_\$})\left[\frac{\kappa^{Q_\$}\theta^{Q_\$}}{\kappa^{Q_\$} - \gamma_1\sigma^{Q_\$}} - \lambda^{Q_\$}(t)\right]dt + \sigma^{Q_\$}\sqrt{\lambda^{Q_\$}(t)}d\widehat{W}^1(t).\end{aligned}$$

Again by using Proposition 11 in Ehlers [17], $\lambda^{Q_€}$ (default intensity under the FSMM measure, $\mathbb{Q}_{€}$) and $\lambda^{Q_\$}$ (default intensity under the measure DSMM, $\mathbb{Q}_{\$}$) satisfy the relationship

$$\lambda^{Q_€} = (1-\delta)\lambda^{Q_\$}. \tag{3.7.12}$$

Then the dynamics of $\lambda^{€}$ is

$$\begin{aligned}d\lambda^{Q_€}(t) &= (1-\delta)(\kappa^{Q_\$} - \gamma_1\sigma^{Q_\$})\left(\frac{\kappa^{Q_\$}\theta^{Q_\$}}{\kappa^{Q_\$} - \gamma_1\sigma^{Q_\$}} - \lambda^{Q_\$}(t)\right)dt + (1-\delta)\sigma^{Q_\$}\sqrt{\lambda^{Q_\$}(t)}d\widehat{W}^1(t),\\ &= (\kappa^{Q_\$} - \gamma_1\sigma^{Q_\$})\left(\frac{(1-\delta)\kappa^{Q_\$}\theta^{Q_\$}}{\kappa^{Q_\$} - \gamma_1\sigma^{Q_\$}} - \lambda^{Q_€}(t)\right)dt + \sqrt{(1-\delta)}\sigma^{Q_\$}\sqrt{\lambda^{Q_€}(t)}d\widehat{W}^1(t),\\ &= \kappa^{Q_€}(\theta^{Q_€} - \lambda^{Q_€}(t))dt + \sigma^{Q_€}\sqrt{\lambda^{Q_€}(t)}d\widehat{W}^1(t),\end{aligned}$$

where

$$\kappa^{Q€} := (\kappa^{Q\$} - \gamma_1 \sigma^{Q\$}), \tag{3.7.13}$$

$$\theta^{Q€} := \frac{(1-\delta)\kappa^{Q\$}\theta^{Q\$}}{\kappa^{Q\$} - \gamma_1 \sigma^{Q\$}}, \tag{3.7.14}$$

$$\sigma^{Q€} := \sqrt{(1-\delta)}\sigma^{Q\$}. \tag{3.7.15}$$

3.8 Econometric Framework and Estimation Results

In this section, we investigate whether it is possible to generate the observed differences in CDS spreads by assuming only correlation between default intensity and exchange rates. Hence we set $\delta = 0$ and $\Phi = 0$. First of all, for the ease of notation, we apply the following change of notation

$$\begin{aligned} \lambda &:= \lambda^{Q\$}, \\ \kappa &:= \kappa^{Q\$}, \\ \theta &:= \theta^{Q\$}, \\ \sigma &:= \sigma^{Q\$}, \\ r_t^{dif} &:= r_t^{\$} - r_t^{€}. \end{aligned}$$

Now to explain our estimation procedure, first, given a set of κ, θ, and σ parameters we extract the λ time series by inverting the pricing function (for the details of pricing formulas we refer the reader to the Appendix 1)

$$CDS_t(M) = f(\lambda_t; \kappa, \theta, \sigma) = 4(1-R)\frac{\int_t^{t+M} E_t^Q \left[\lambda_u^Q \exp(-\int_t^u (r_s + \lambda_s^Q) ds)\right] du}{\sum_{i=1}^{4M} E_t^Q \left[\exp(-\int_t^{t+0.25i} (r_s + \lambda_s^Q) ds)\right]}. \tag{3.8.1}$$

More explicitly, given a set of κ, θ, and σ parameters, we use the 5-year CDS spreads denominated in USD to recover a time series for λ by means of a non-linear optimization algorithm:

$$\lambda_t = f^{-1}(CDS_t^{\$}(5); \kappa, \theta, \sigma). \tag{3.8.2}$$

Then we compute likelihood function to maximize by using the joint Euler approximation of the exchange rate and default intensity processes as the following: First, we rewrite the exchange rate and default intensity dynamics as

$$\begin{aligned} d\lambda(t) &= (\kappa - \phi_1\sigma)\left[\frac{\kappa\theta}{\kappa - \phi_1\sigma} - \lambda(t)\right]dt + \sigma\sqrt{\lambda(t)}d\widetilde{W}^1(t), & (3.8.3)\\ \frac{dX_t}{X_t} &= (r_t^{dif} + \gamma_1\phi_1\lambda_t + \gamma_2\phi_2)dt + \gamma_1\sqrt{\lambda_t}d\widetilde{W}_t^1 + \gamma_2 d\widetilde{W}_t^2. & (3.8.4) \end{aligned}$$

By defining

$$Y_t = \begin{bmatrix} \lambda_t \\ X_t \end{bmatrix}$$

we have

$$dY_t = \underbrace{\begin{bmatrix} \kappa\theta - (\kappa - \phi_1\sigma)\lambda \\ X_t(r_t^{dif} + \gamma_1\phi_1\lambda_t + \gamma_2\phi_2) \end{bmatrix}}_{b(t,Y_t)} dt + \underbrace{\begin{bmatrix} \sigma\sqrt{\lambda_t} & 0 \\ X_t\gamma_1\sqrt{\lambda_t} & X_t\gamma_2 \end{bmatrix}}_{\sigma(t,Y_t)} \underbrace{\begin{bmatrix} d\widetilde{W}_t^1 \\ d\widetilde{W}_t^2 \end{bmatrix}}_{d\widetilde{W}_t}. \tag{3.8.5}$$

Hence we obtain

$$dY_t = b(t, Y_t)dt + \sigma(t, Y_t)d\widetilde{W}_t \quad \text{with} \quad Y_0 = y_0 \in \mathbb{R}^2. \tag{3.8.6}$$

Let $t_k = k\Delta, k = 0, 1, ..., N$ be the time discretization steps with $\Delta > 0$. Consider the Euler scheme of the process $(Y_t)_{t\geq 0}$ starting from $Y_0 = y_0$.

$$Y_{t_{k+1}} = Y_{t_k} + b(t_k, Y_{t_k})\Delta + \sigma(t_k, Y_{t_k})(\widetilde{W}_{t_{k+1}} - \widetilde{W}_{t_k}). \tag{3.8.7}$$

Define the functions

$$\mu_k(y) := y + b(t_k, y)\Delta, \quad (3.8.8)$$

$$\Sigma_k(y) := \sigma(t_k, y)\sigma(t_k, y)^\intercal\Delta. \quad (3.8.9)$$

It is clear that

$$Y_{t_{k+1}} \backslash Y_{t_k} \sim N(\mu_k(Y_{t_k}), \Sigma_k(Y_{t_k})), \quad (3.8.10)$$

and let's denote its density function by $\Phi_{\mu_k(y),\Sigma_k(y)}(.)$.

$$\Phi_{\mu_k(y),\Sigma_k(y)}(Y_{t_{k+1}}) = \frac{1}{2\pi\sqrt{|\Sigma_k(Y_{t_k})|}} exp\left(-\frac{1}{2}\Xi_k\right) \quad (3.8.11)$$

where $|\Sigma_k(Y_{t_k})|$ denotes the determinant of $\Sigma_k(Y_{t_k})$ and

$$\Xi_k := \left(Y_{t_{k+1}} - \mu_k(Y_{t_k})\right)^\intercal \Sigma_k^{-1} \left(Y_{t_{k+1}} - \mu_k(Y_{t_k})\right).$$

Let $\Theta = (\kappa, \theta, \sigma, \phi_1, \phi_2, \gamma_1, \gamma_2)$. Then the likelihood function can be written as

$$\mathcal{L}(\Theta|Y_{t_1}, Y_{t_2}, ..., Y_{t_N}) = \prod_{k=0}^{N-1} \Phi_{\mu_k(y),\Sigma_k(y)}(Y_{t_{k+1}}|\Theta), \quad (3.8.12)$$

$$= \frac{1}{(2\pi)^N} \prod_{k=0}^{N-1} \frac{1}{\sqrt{|\Sigma_k(Y_{t_k})|}} exp\left(-\frac{1}{2}\Xi_k\right). \quad (3.8.13)$$

In practice, it is often more convenient to work with the logarithm of the likelihood function, called the log-likelihood,

$$ln\mathcal{L}(\Theta|Y_{t_1}, Y_{t_2}, ..., Y_{t_N}) = \sum_{k=0}^{N-1} ln\Phi_{\mu_k(y),\Sigma_k(y)}(Y_{t_{k+1}}|\Theta), \quad (3.8.14)$$

$$= -Nln(2\pi) - \frac{1}{2}\sum_{k=0}^{N-1} ln(|\Sigma_k(Y_{t_k})|) - \frac{1}{2}\sum_{k=0}^{N-1} \Xi_k \quad (3.8.15)$$

To estimate the model parameters and CDS spreads, we take the dollar interest rate to be constant at 1 %, $r_t^{\$} = 0.01$, and euro interest rate to be constant at 2 %, $r_t^{€} = 0.02$ (There is theoretical and empirical evidence that the interest rates have small effect on CDS rates hence taking interest rates to be constant does not affect our results significantly). Our CDS database provides the recovery rates for each sovereign for the entire period time. For the countries under investigation, we also take the recovery rates to be constant at 40 %, $R = 0.40$.

For the estimation of the initial parameters, we first use the approximation $CDS_t(5) \approx (1 - R)\lambda_t$ to extract the default intensities from USD CDS spreads. Then we estimate $\kappa^P, \theta^P, \sigma^P$ by running a linear regression on the discrete version of the equation (3.7.6) and then we use the equations following (3.7.6) to find the initial parameter estimates for $\kappa^{Q_\$}, \theta^{Q_\$}, \sigma^{Q_\$}$. Similarly, we run a linear regression on the exchange rate process (3.8.4) and on a function of it, X_t^2, to estimate the initial values for $\phi_1, \phi_2, \gamma_1, \gamma_2$. However, our empirical investigation reveals that assigning zero for the initial values of ϕ_1, ϕ_2 yields better parameters estimates.

After finding the initial values, we maximize the log-likelihood function by exactly matching USD CDS spreads at each iteration. As it can be observed from Figures (3.4)-(3.6) the difference between model and market USD CDS rates is not different from zero for each sovereign as expected. We use the equation (3.7.13) to find $\kappa^{Q_€}, \theta^{Q_€}, \sigma^{Q_€}$. The list of estimated parameters is given in the Table (3.5) . The table shows that the market prices of risk values ϕ_1, ϕ_2 are very close to zero for each country. We also observe that the sign of the γ_1 which is the sign of the instantaneous correlation between default intensity and exchange rates is negative for all countries as expected.

As Figures (3.4)-(3.6) provide clear evidence, assuming only correlation between default intensity and exchange rates, the resulting spread between USD and EUR CDS rates was in no case even close to the spread empirically observed in the data. Hence our results also verify that the correlation by itself is not enough to explain real market data. Following Ehlers [17], consequently, we also reject the pure diffusion model hence there must be jumps in the exchange rate at default if the model were correct.

Table 3.5: Parameter Estimates
This table shows ML estimates of risk neutral parameters under both domestic and foreign risk neutral measures ($r_t^{\$} = 0.01$, $r_t^{€} = 0.02$, and $R = 0.40$ in our computations).

	Belgium	**France**	**Germany**	**Ireland**	**Italy**	**Spain**
$\kappa^{Q_\$}$	0.258	0.139	0.170	0.062	0.180	0.254
$\theta^{Q_\$}$	0.032	0.027	0.010	0.163	0.052	0.039
$\sigma^{Q_\$}$	0.128	0.086	0.060	0.142	0.137	0.141
$\kappa^{Q_€}$	0.271	0.184	0.206	0.074	0.186	0.260
$\theta^{Q_€}$	0.030	0.020	0.009	0.136	0.050	0.038
ϕ_1	0.002	-0.012	-0.001	-0.001	0.000	0.001
ϕ_2	-0.005	0.003	-0.005	0.004	-0.002	0.001
γ_1	-0.104	-0.519	-0.603	-0.088	-0.046	-0.044
γ_2	0.137	0.122	0.126	0.097	0.153	0.125

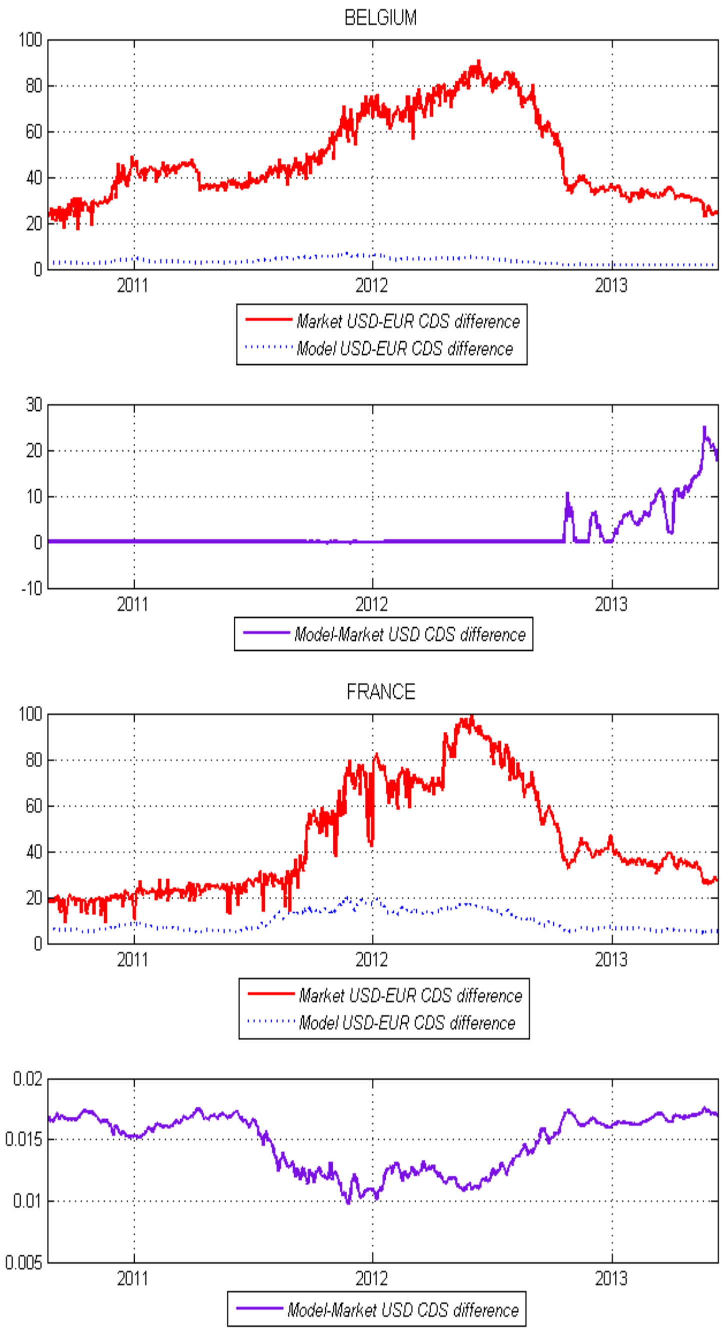

Figure 3.4: *Market and model CDS spreads* The figures above display the observed USD and EUR CDS differences in the market, the differences computed with our stochastic model and also the model and market USD CDS spread differentials for the countries Belgium and France for the period 20 August 2010 to 12 June 2013.

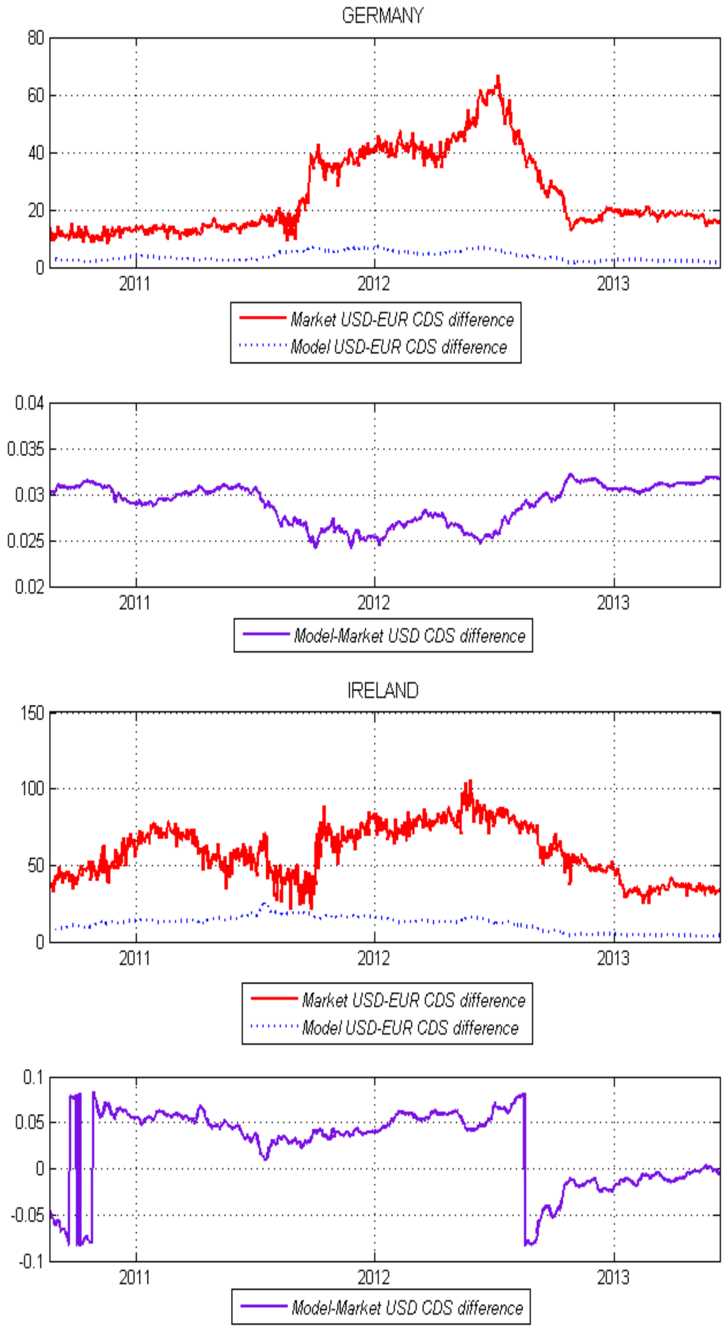

Figure 3.5: *Market and model CDS spreads* The figures above display the observed USD and EUR CDS differences in the market, the differences computed with our stochastic model and also the model and market USD CDS spread differentials for the countries Germany and Ireland for the period 20 August 2010 to 12 June 2013.

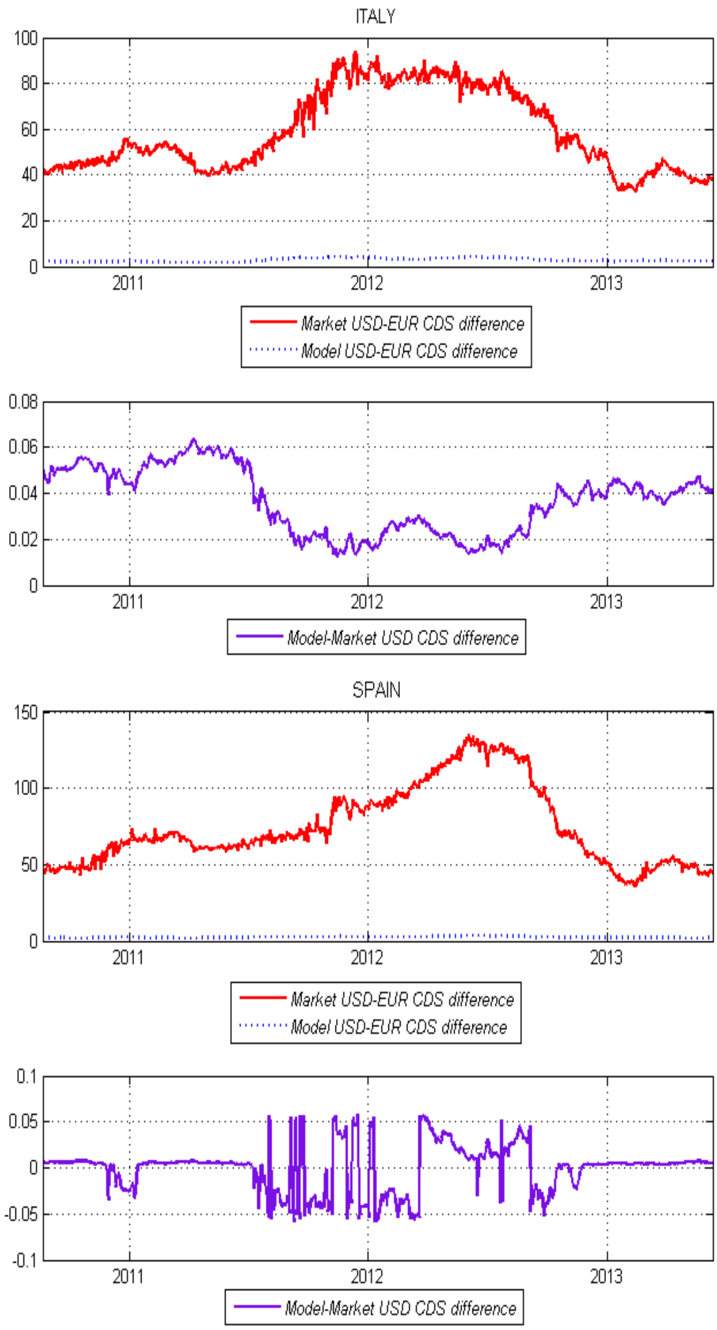

Figure 3.6: *Market and model CDS spreads* The figures above display the observed USD and EUR CDS differences in the market, the differences computed with our stochastic model and also the model and market USD CDS spread differentials for the countries Italy and Spain for the period 20 August 2010 to 12 June 2013.

3.9 Conclusion

The aim of this paper is to study dual currency credit default swaps from both theoretical and empirical perspectives. For this purpose, we study six European markets that quote CDS both in USD and EUR: Belgium, France, Germany, Ireland, Italy, and Spain for the period August 20, 2010 to June 13, 2013. On the empirical side, we fist calculate a proxy of deviation from the LOP called Quanto CDS and we observe that Quanto CDSs are considerably large and state-dependent for each sovereign. Our second finding is related to the price discovery in the corresponding CDSs, which sheds light to the natural habitat of traders. We note that dollar-denominated CDSs always tend to lead the price variations of EUR-denominated CDSs for all countries. We also conduct a factor analysis which is useful in reducing the dimension by concentrating on a few important factors that represent the main sources of variation in the dual currency CDS market. We find that the first two factors statically significant during the entire period of the time. On the theoretical side, we consider a correlation based diffusion model with a jump component to test the currency dependence of credit spreads. Our results, being consistent with the literature, verify that the correlation by itself is not enough to explain observed spread differences between USD and EUR CDS spreads.

Appendix 1: Analytical Solution for CDS prices

This section describes the pricing of the default swap as in Longstaff et al. [44]. Assume that the default intensity follows the CIR process with the following parameters

$$d\lambda^Q(t) = \kappa^Q(\theta^Q - \lambda^Q(t))dt + \sigma^Q\sqrt{\lambda^Q(t)}dW(t).$$

Assuming that the interest rate process is independent to the intensity one, the CDS price with quarterly premia is expressed as follows

$$\begin{aligned} CDS_t(M) &= 4(1-R)\frac{\int_t^{t+M} E_t^Q\left[\lambda_u^Q \exp(-\int_t^u (r_s+\lambda_s^Q)ds)\right]du}{\sum_{i=1}^{4M} E_t^Q\left[\exp(-\int_t^{t+0.25i}(r_s+\lambda_s^Q)ds)\right]} \\ &= 4(1-R)\frac{\int_t^{t+M} e^{B(s)\lambda_t} D(s)(G(s)+H(s)\lambda_t)ds}{\sum_{i=1}^{4M} A(t+0.25i)e^{B(t+0.25i)\lambda_t}D(t+0.25i)} \end{aligned} \tag{3.9.1}$$

where

$$\begin{aligned} A(s) &= \exp\left(\frac{\kappa^Q\theta^Q(\kappa^Q+w)}{(\sigma^Q)^2}s\right)\left(\frac{1-v}{1-v\exp(ws)}\right)^{\frac{2\kappa^Q\theta^Q}{(\sigma^Q)^2}} \\ B(s) &= \frac{\kappa^Q - w}{(\sigma^Q)^2} + \frac{2w}{(\sigma^Q)^2(1-v\exp(ws))} \\ G(s) &= \frac{\kappa^Q\theta^Q}{w}(\exp(ws)-1)\exp\left(\frac{\kappa^Q\theta^Q(\kappa^Q+w)}{(\sigma^Q)^2}s\right)\left(\frac{1-v}{1-v\exp(ws)}\right)^{\frac{2\kappa^Q\theta^Q}{(\sigma^Q)^2}+1} \end{aligned}$$

$$\begin{aligned} H(s) &= \exp\left(\frac{\kappa^Q\theta^Q(\kappa^Q+w)+w(\sigma^Q)^2}{(\sigma^Q)^2}s\right)\left(\frac{1-v}{1-v\exp(ws)}\right)^{\frac{2\kappa^Q\theta^Q}{(\sigma^Q)^2}+2} \\ D(s) &= \exp\left(\int_0^s r_u du\right) \\ w &= \sqrt{2\sigma^2+(\kappa^Q)^2} \\ v &= \frac{\kappa^Q+w}{\kappa^Q-w} \end{aligned}$$

Bibliography

[1] Lustig, H., Roussanov, N., and Verdelhan, A. (2011). Common risk factors in currency markets. *Review of Financial Studies*, 24(11) 3731-3777.

[2] Breger, L. (2003). Fixed Income Risk Modelling for Portfolio Managers *Review of Financial Studies*, 24(11) 3731-3777.

[3] Lustig, H. and Verdelhan, A. (2007). The cross section of foreign currency risk premia and consumption growth risk. *Technical Report, Fixed Income Research, Barra, Inc.*

[4] Eun, C. S. and Sabherwal, S. (2003). Cross-Border Listings and Price Discovery: Evidence from US-Listed Canadian Stocks. *The Journal of Finance*, 58(2), 549–576.

[5] Gonzalo, J. and Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics*, 13(1), 27–35.

[6] Scherer, K. P. and Avellaneda, M. (2002). All for one... one for all? A principal component analysis of Latin American brady bond debt from 1994 to 2000. *International Journal of Theoretical and Applied Finance*, 5(1), 79–106.

[7] Vassalou, M. (2000). Exchange rate and foreign inflation risk premiums in global equity returns. *Journal of International Money and Finance*, 19(3), 433–470.

[8] De Santis, G. and Gerard, B. (1998). How big is the premium for currency risk?. *Journal of Financial Economics*, 49(3), 375–412.

[9] Domowitz, I., Glen, J., Madhavan, A. (1998). Country and Currency Risk Premia in an Emerging Market. *Journal of Financial and Quantitative Analysis*, 33(2), 189-216.

[10] Dumas, B. and Solnik, B. (1995). The world price of foreign exchange risk. *The journal of finance*, 50(2), 445–479.

[11] Pukthuanthong, K. and Roll, R. (2009). Global market integration: An alternative measure and its application. *Journal of Financial Economics*, 94(2), 214–232.

[12] De Jong, F. and de Roon, F. A. (2005). Time-varying market integration and expected returns in emerging markets. *Journal of Financial Economics*, 78(3), 583–613.

[13] Bekaert, G. and Harvey, C. R. (1995).Time-varying world market integration. *The Journal of Finance*, 50(2) 403–444.

[14] Stulz, R. M. (1981). A model of international asset pricing. *Journal of Financial Economics*, 9(4) 383–406.

[15] Solnik, B. H. (1974). Why not diversify internationally rather than domestically?. *Financial analysts journal*, 48–54.

[16] Adler, M., and Dumas, B. (1983). International portfolio choice and corporation finance: A synthesis. *The Journal of Finance*, 38(3), 925-984.

[17] Ehlers, P. (2007). Pricing credit derivatives. *Ph.D. Thesis* Swiss Federal Institute of Technology.

[18] Artzner, P., Delbaen, F. (1995). Default Risk Insurance and Incomplete Markets. *Mathematical Finance*, 5, 187-195.

[19] Black, F., Cox, J. C. (1976). Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. *Journal of Finance*, 31(2), 351-367.

[20] Blanco, R., Brennan, S., and Marsh, I. W. (2005). An Empirical Analysis of the Dynamic Relation between Investment-Grade Bonds and Credit Default Swaps. *Journal of Finance*, 60(5), 2255-2281.

[21] Brunnermeier, M. K., Pedersen, L. H. (2009). Market Liquidity and Funding Liquidity. *Review of Financial Studies*, 22(6), 2201-2238.

[22] Buraschi, A., Sener, E., Menguturk, M. (2013). The Dynamics of Limits to Arbitrage: An Empirical Investigation. Working Paper

[23] Buraschi, A., Trojani, F., and Vedolin, A. (2009). Economic Uncertainty, Differences in Beliefs and Credit Markets. *Journal of Political Economy*, under submission.

[24] Cariboni, J., Schoutens, W. (2007). Pricing Credit Default Swaps under Levy Models. *Journal of Computational Finance*, 10(4), 1-21.

[25] Chen, L., Lesmond, and D. A., Wei, J. (2007). Bond Liquidity Estimation and the Liquidity Effect in Yield Spreads.*Journal of Finance*, 62, 119-49.

[26] Collin-Dufresne, P., Goldstein, R. S., and Martin, J. S. (2001). The Determinants of Credit Spread Changes. *Journal of Finance*, 56(6), 2177-2207.

[27] Duffie, D., Singleton, K. J. (1999). Modeling Term Structures of Defaultable Bonds. *Review of Financial Studies*, 12(4), 687-720.

[28] Driessen, J. (2005). Is Default Event Risk Priced in Corporate Bonds? *Review of Financial Studies*, 18(1), 165-195.

[29] Duffee, G. R. (1999). Estimating the Price of Default Risk. *Review of Financial Studies*, 12(1), 197-226.

[30] Elliott, R., Jeanblanc M., and Yor, M. (2000). On Models of Default Risk. *Mathematical Finance*, 10, 179-195.

[31] Elton, E. J., Gruber, M. J., Agrawal, D., and Mann, C. (2001). Explaining the Rate Spread on Corporate Bonds. *Journal of Finance*, 56(1), 247-277.

[32] Ehlers, P., Schönbucher, P. (2004). The Influence of FX Risk on Credit Spreads. *National Centre of Competence in Research Financial Valuation and Risk Management Working Paper Series* No. 214.

[33] Ericsson, J., Renault, O. (2006). Liquidity and Credit Risk. *Journal of Finance*, 61(5), 2219-2250.

[34] Fitch Soutions. (2010, 30 March). Greece Sovereign CDS - A History of Myth and Reality. Fitch Solutions Special Report: Aubrey, T., Brigo, D.

[35] Feldhatter, P., Lando, D. (2008). Decomposing Swap Spreads. *Journal of Financial Economics*, 88, 375-405.

[36] Hasbrouck, J. (1995). One Security, Many Markets: Determining the Contributions to Price Discovery. *Journal of Finance*, 50(4), 1175-1199.

[37] Hull, J., Predescu, M., and White, A. (2004). The Relationship Between Credit Default Swap Spreads, Bond Yields, and Credit Rating Announcements. *Journal of Banking and Finance*, 28, 2789-2811.

[38] Jankowitsch, R., Pichler, S. (2005). Currency Dependence of Corporate Credit Spreads. *Journal of Risk*, 8, 1-24.

[39] Jarrow, R., Turnbull, S. (1995). Pricing Derivatives on Financial Securities Subject to Credit Risk. *Journal of Finance*, 50(1), 53-85.

[40] Johansen, S. (1988). Statistical Analysis of Cointegrating Vectors. *Journal of Economic Dynamics and Control*, 12, 231-254.

[41] Kercheval, A., Goldberg, L., Breger, L. (2003). Modeling Credit Risk. *Journal of Portfolio Management*, 29, 90-100.

[42] Landschoot, A. V. (2008). Determinants of Yield Spread Dynamics: Euro versus US Dollar Corporate Bonds. *Journal of Banking and Finance*, 32, 2597-2605.

[43] Leland, E.H., 1994. Corporate Debt Value, Bond Covenants, and Optimal Capital Structure. *Journal of Finance*, 49(4), 1213-1252.

[44] Longstaff, F. A, Neis, E., Mithal, S. (2005). Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit-Default Swap Market. *Journal of Finance*, 60, 2213-2254.

[45] Longstaff, F. A, Pan, J., Pedersen, L. H., and Singleton, K. J. (2011). How Sovereign is Sovereign Credit Risk? NBER Working Paper Series No: W13658.

[46] Merrill Lynch (2000). Merrill Lynch Global Bond Indices-Rules of Construction and Calculation Methodology.

[47] Merton, R. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, 29, 789-821.

[48] Pan, J. and Singleton, K. J. (2008). Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads. *Journal of Finance*, 63(5), 2345-2384.

[49] Sato, K. (2000). Levy Processes and Infinitely Divisible Distributions, United Kingdom, Cambridge Studies in Advanced Mathematics, Cambridge University Press.

[50] Zhou, C. (2001). The Term Structure of Credit Spreads with Jump Risk. *Journal of Banking and Finance*, 25(11), 2015-2040.

Chapter 4

Partial Hedging and Cash Requirement in Discrete Time

Erdinç Akyıldırım

4.1 Introduction

One of the most prominent problems in financial mathematics is the correct pricing and hedging of financial instruments also known as derivatives. The existence of a unique Equivalent Martingale Measure (EMM) in the framework of complete market models which are arbitrage-free makes this problem straightforward. The pricing of contingent claims is achieved by replicating the payoff of a derivative security by a self-financing trading strategy. In this kind of strategy, the underlying assets are traded without adding or withdrawing money from the total budget. In complete markets, every contingent claim is attainable i.e. each derivative product can be exactly replicated by a self-financing trading strategy. Therefore, the cost of the replication is equal to the unique arbitrage free price of the derivative and it can be computed as the expectation of the claim under the unique EMM.

Black and Scholes [7] and Merton [18] are the most well-known examples of complete market models. Rubinstein [9] states that "the Black & Scholes model is widely viewed as one of the most successful in the social sciences and perhaps, including its binomial extension, the most widely used formula, with embedded probabilities, in human history".

However, in the Black-Scholes-Merton world there are some simplifying assumptions such as constant volatility and mean, no transaction costs, continuous trading, and no taxes. On the other hand, empirical studies show that this restrictive set of assumptions are not consistent with real world data. Hence relaxing any one of these assumptions leads to an incomplete market or a market with friction. In these markets financial instruments carry an intrinsic risk which cannot be perfectly hedged. Therefore not every contingent claim is attainable in such markets. In this case, it is well known that, the EMM is no longer unique and there is set of EMMs under which the discounted prices of the instruments are martingales. Thus, there exist infinitely many arbitrage free price processes for a given derivative. An analogous difficulty arises when there are market frictions such as portfolio constraints, transaction costs, taxes or liquidity issues.

In incomplete markets or markets with friction we can only determine no arbitrage intervals, outside of which there is arbitrage. One is then faced with the issue of choosing an appropriate price among many. At this point, the choice depends crucially on the risk preferences of the individual investors and the real world measure. It is important to note that in a complete market the risk preferences do not play any role because, at least in theory, perfect hedging completely removes the risk associated with hedging as well as the opportunity to make a profit. However, in the other case, option pricing has to be based on individual's attitude toward risk which is modelled by a utility or loss function. For instance, a fully risk averse investor may choose to stay on the safe side by using a super-hedging strategy. The idea of super-hedging has been introduced to the literature by Bensaid and Lesne [1] in discrete time and El Karoui and Quenez [12] in continuous time. The aim in super-hedging is to generate a final wealth that dominates the payoff of the contingent claim. It is proved in [12] that the value of the cheapest portfolio to dominate the pay-off of a contingent claim is the same as the supremum of the expected values of the pay-off of the claim over all EMMs. A common criticism about super-hedging is that in some situations the cost of super-hedging can be too high from a practical point of view.

On the other hand if the investor is willing to take some risk, then there should be a reduction in the initial cost as a risk premium. Similarly, if the investor wants to invest less capital than the perfect or super-hedging price of the liability, then some shortfall risk has to be accepted. Föllmer and Leukert [10] use probability as a risk measure to quantify this shortfall risk. In their seminal paper [10], quantile hedging is described as the optimal hedge when the initial capital is less than the minimal super-hedging or perfect hedging cost. In particular, they determine the minimal amount of initial capital an investor can save by accepting a certain shortfall probability. Equivalently, they find the maximal probability of a successful hedge the investor can achieve if she is unwilling to put up the initial amount of capital required by a super-hedging (or a replication) strategy. In a complete market, by using the Neyman-Pearson Lemma in mathematical statistics, they show that quantile hedging is simply perfect hedging of a cheaper claim whose Black-Scholes price is equal to the given initial premium.

Cvitanic and Spivak [16] and the references therein also studied the quantile heding and related problems in continuous time. The solution in complete markets was re-derived by [16] using a duality approach in the context of utility maximization. Then this approach was modified to solve the problem in a market with partial information and in markets in which the wealth process of the agent has a non-linear drift. Krutchenko and Melnikov [19] extend the work of [10] to a jump-diffusion financial market model. Lindberg [17] considers the quantile hedging problem as a knapsack problem which is a widely researched subject in linear programming. He develops an efficient algorithm which works for European options in a complete discrete-time market model. Perez-Hernandez [21] studies the quantile hedging problem for American contingent claims in an infinite-state space setting from the perspective of the writer of the claim and Pınar [22] investigates the same problem from the point of view of a buyer of a contingent claim by minimizing the expected failure ratio. There are also a number of papers applying the quantile hedging approach to the insurance setting. (see Melnikov and Skornyakova [4] and the references therein).

In a recent paper,Bouchard et al. [11] provide a different approach to quantile hedging. They consider the more general problem of finding the minimal initial data of a controlled process which guarantees reaching a controlled target with a given probability of success. As a special case, they focus on the quantile hedging problem and reproduce the explicit solution of [10] in continuous time. In this article one of our goals is to further understand and develop their techniques in discrete time. We believe that our discrete time model has the advantage of streamlining the main ideas in [11] and bringing numerical difficulties to the surface. Our discrete time model can also recover the solutions for utility indifference pricing, good deal pricing, and expected shortfall, but our main contribution to the literature is in the context of quantile hedging.

As mentioned above there are a number of papers which study the quantile hedging problem in various frameworks. However they require vastly different constructions under different market conditions. The main advantage of our discrete time model is that we can handle different market structures such as exotic options and markets with portfolio constraints by only slightly modifying our original method. The detailed numerical analysis of the quantile hedging problem in these varying frameworks and its interpretation is provided in our numerics section.

The paper is organized as follows: The next section presents a description of our financial market model. The general problem and its equivalent formulation are studied in Section 3. In Section 4, we derive the dynamic programming principle and illustrate the efficient algorithm for the application of dynamic programming. Section 5 includes application of our method in the context of the quantile hedging. Numerical results are also provided in Section 5. The paper is concluded in Section 6.

4.2 Complete Financial Market Setting

We consider a discrete-time complete market framework with one risk-free and one risky asset for a continuous-time stochastic target problem. We fix a time horizon, or equivalently a maturity, $T > 0$ and a time discretization $h := \frac{T}{n}$ with a large integer n. We assume that the risky asset price process X_t evolves according to

$$X_{t+1} = X_t \left(1 + \mu h + \xi_{t+1}\sigma\sqrt{h}\right), \tag{4.2.1}$$

where μ is the mean return rate, σ is the volatility, the noise $\{\xi_t\}$ is an i.i.d. random sequence with

$$\mathrm{E}\left[\xi_t\right] = 0,\ \mathrm{E}\left[\ \xi_t^2\right] = 1.$$

We assume that ξ_t takes values in $\{-1,+1\}$ and hence we have a binomial tree structure for the risky asset. Clearly, the no-arbitrage condition yields

$$1+\mu h-\sigma\sqrt{h}<1+r<1+\mu h+\sigma\sqrt{h}, \tag{4.2.2}$$

which will also be useful to speed up the computations in our model. For $0\le t\le s\le T$, we set the filtration

$$\mathcal{F}_t^s := \sigma(\xi_t,\dots,\xi_s)=\sigma(X_t,\dots,X_s). \tag{4.2.3}$$

to be generated by the spot price process X_t. Hence the sigma-algebra $\mathcal{F}_t^s$ contains the information from time t to s. We take the interest rate $r=0$ for the simplicity of the computations. Under the usual self-financing assumption, the *wealth process* Y_t follows the dynamics

$$Y_{t+1}=Y_t+Z_t(X_{t+1}-X_t), \tag{4.2.4}$$

where Z_t is the *portfolio process* and $\mathcal{A}^t$ is the set of all $\mathcal{F}$- adapted, zero- admissible portfolio (control) processes $\{Z.\}$ with $Z_t=z$. Given initial conditions (t,x) and (t,y) we employ the notation $X_\cdot^{t,x}$, $Y_\cdot^{t,y,Z.}$ i.e. spot price and wealth processes satisfy the initial conditions

$$X_t^{t,x}=x,\quad Y_t^{t,y,Z}=y.$$

The *liability* is assumed to be Markov with nonnegative payoff $g(X_T)$.

4.3 The general problem and an equivalent formulation

Given a strictly increasing and concave utility function U and a *threshold* utility level u^*, we consider the problem of finding the minimal initial value of a controlled wealth process such that the expected utility from the final wealth minus liability is guaranteed to reach a pre-specified target utility level u^*

$$V(t,x):=\inf\left\{\, y\ :\ \exists\{Z\}\in\mathcal{A}^t \text{ s.t. } \mathrm{E}\left[U\left(Y_T^{t,y,Z}-g(X_T^{t,x})\right)\right]\ge u^*\,\right\}. \tag{4.3.1}$$

Given an admissible portfolio process Z, define

$$u_s:=\mathrm{E}\left[U(Y_T^{t,y,Z}-g(X_T^{t,x}))\mid \mathcal{F}_t^s\,\right]. \tag{4.3.2}$$

It is clear that u_s depends on the initial conditions, as well as on the portfolio process but from now on we will suppress these dependence. First we note that

$$\mathrm{E}\left[u_{s+1} \mid \mathcal{F}_t^s\right] \;=\; \mathrm{E}\left[U\left(Y_T^{t,y,Z} - g(X_T^{t,x})\right) \mid \mathcal{F}_t^{s+1} \mid \mathcal{F}_t^s\right] = u_s.$$

Hence u_s is a $\mathcal{F}_t^s$ martingale. Then by Martingale Representation Theorem, there exists an adapted α_s such that

$$u_{s+1} = u_s + \alpha_s \xi_{s+1}. \tag{4.3.3}$$

Moreover, in view of our original problem (4.3.1)

$$u_t = \mathrm{E}\left[U\left(Y_T^{t,y,Z} - g(X_T^{t,x})\right)\right] \geq u^*.$$

We expect this inequality to saturate. Therefore, we use the initial condition

$$u_t = u^*. \tag{4.3.4}$$

We use the notation $u_\cdot^{t,u^*,\alpha}$. Since

$$u_T^{t,u^*,\alpha} = U\left(Y_T^{t,y,Z} - g(X_T^{t,x})\right) \;\Leftrightarrow\; Y_T^{t,y,Z} = U^{-1}(u_T^{t,u^*,\alpha}) + g(X_T^{t,x}),$$

we formulate the following *super-replication* problem

$$v(t,x,u^*) \;:=\; \inf\left\{\, y \;:\; \exists\{Z\} \in \mathcal{A}^t,\ \{\alpha.\} \text{ s.t. } Y_T^{t,y,Z} \geq U^{-1}(u_T^{t,u^*,\alpha}) + g(X_T^{t,x}) \ a.s. \,\right\}. \tag{4.3.5}$$

Theorem 4.3.1. If $g(X_T)$ is a European contingent claim then $V(t,x) = v(t,x,u^*)$.

Proof. Let $Z^\epsilon \in \mathcal{A}_t^s$ be a portfolio process satisfying

$$\mathrm{E}\left[U\left(Y_T^{t,y,Z} - g(X_T^{t,x})\right)\right] \geq u^* \ \text{ with } \ y = V(t,x) + \epsilon.$$

Let u_s^ϵ be as in (4.3.2). Then, clearly

$$u_T^\epsilon \;=\; U\left(Y_T^{t,y,Z} - g(X_T^{t,x})\right).$$

Thus, Z^ϵ with initial wealth $V(t,x)+\epsilon$ satisfies the constraint for problem (4.3.5) and by definition of v we have $v(t,x,u^*) \le V+\epsilon$.

To prove the reverse inequality let $Z^\epsilon \in \mathcal{A}_t^s$ and α^ϵ be super-replicating for the problem (4.3.5) with initial wealth $y = v(t,x,u^*)+\epsilon$. Then,

$$
\begin{aligned}
Y_T^{t,y,Z} &\ge U^{-1}(u_T^{t,u^*,\alpha}) + g(X_T^{t,x}),\\
&\Rightarrow U\left(Y_T^{t,y,Z} - g(X_T^{t,x})\right) \ge u_T^{t,u^*,\alpha},\\
&\Rightarrow \mathrm{E}\left[U\left(Y_T^{t,y,Z} - g(X_T^{t,x})\right)\right] \ge \mathrm{E}\left[u_T^{t,u^*,\alpha}\right] = u^*.
\end{aligned}
$$

Hence, Z^ϵ with initial wealth $v(t,x,u^*)+\epsilon$ satisfies the constraint for the problem (4.3.1) and we have $V(t,x) \le v(t,x,u^*)+\epsilon$ by definition of $V(t,x)$.

4.4 Dynamic Programming

In mathematical finance, the classical super-hedging problem, in general, is solved by the dual formulation approach which assumes that the wealth dynamics is linear in the control and that the stocks prices are not influenced by the trading strategy. Hence more general dynamics or constraints such as gamma constraints cannot be treated with this approach. Motivated by this problem, Soner and Touzi [15] introduce stochastic target techniques to provide a PDE characterization of the super-hedging price of an European claim under gamma constraints. Their main contribution is a dynamic programming principle which is directly written on the associated stochastic target problem. Bouchard et al. [11] extend the work of [15] from super-hedging problems to the pricing problems under risk constraints such as quantile hedging. Hence the following dynamic programming results can easily be proved by standard techniques in Fleming and Soner [20] and Soner and Touzi [15].

For any stopping time, $t < \tau \le T$, the minimum super-replication cost $v(t,x,u)$ satisfies

$$
v(t,x,u) \ := \ \inf\left\{ \ y \ : \ \exists\{Z\} \in \mathcal{A}^t, \ \{\alpha.\}, \text{ s.t. } Y_\tau^{t,y,Z} \ge v(\tau, X_\tau^{t,x}, u_\tau^{t,u,\alpha}) \ a.s. \ \right\} \tag{4.4.1}
$$

In discrete time, if we take $\tau = t+1$ then we can conclude that

$v(t,x,u) =$ min y among all y satisfying

$$Y_{t+1}^{t,y,Z} \geq v(t+1, X_{t+1}^{t,x}, u_{t+1}^{t,u,\alpha}),$$

both when the risky asset moves up and down. Using the wealth equation (4.2.4), we rewrite the above equation as

$$y + Z_t(X_{t+1} - X_t) \geq v(t+1, X_{t+1}^{t,x}, u_{t+1}^{t,u,\alpha}).$$

More explicitly, if we define

$$x_{up} := x\left(1 + \mu h + \sigma\sqrt{h}\right), \tag{4.4.2}$$

$$x_{down} := x\left(1 + \mu h - \sigma\sqrt{h}\right), \tag{4.4.3}$$

then our dynamic programming equation looks like

$$\begin{aligned} v(t,x,u) &= \min y \quad \text{among all y satisfying} \\ &\quad y + z\left(x_{up} - x\right) \geq v\left(t+1, x_{up}, u+\alpha\right), \\ &\quad y + z\left(x_{down} - x\right) \geq v\left(t+1, x_{down}, u-\alpha\right). \end{aligned}$$

Hence,

$$v(t,x,u) = \inf_{\alpha,z}\left\{\max\left\{\begin{array}{l} v\left(t+1, x_{up}, u+\alpha\right) - z\left(x_{up} - x\right), \\ \\ v\left(t+1, x_{down}, u-\alpha\right) - z\left(x_{down} - x\right) \end{array}\right\}\right\}.$$

The above equation is complemented by the terminal condition

$$v(T,x,u) = U^{-1}(u) + g(x). \tag{4.4.4}$$

Lemma 4.4.1. Let $\mathcal{A} \subseteq \mathbb{R}$ and $F, G : \mathcal{A} \times \mathbb{R} \to \mathbb{R}$ be two functions such that for given α, $F(\alpha, z)$ is a non-increasing and $G(\alpha, z)$ is a non-decreasing function of z. Assume that for any α, there exists at least one $z^*(\alpha)$ such that $F(\alpha, z^*(\alpha)) = G(\alpha, z^*(\alpha))$ then the following holds

$$\inf_z \{\max\{F(\alpha, z), G(\alpha, z)\}\} = F(\alpha, z^*(\alpha)) = G(\alpha, z^*(\alpha)). \tag{4.4.5}$$

Proof. For any $\alpha \in A \subseteq \mathbb{R}$ given, it is clear that

$$F(\alpha, z^*(\alpha)) = G(\alpha, z^*(\alpha)) \geq \inf_z \{\max\{F(\alpha, z), G(\alpha, z)\}\},$$

where $z^*(\alpha)$ is defined as $F(\alpha, z^*(\alpha)) = G(\alpha, z^*(\alpha))$. To show the other side of the inequality, assume for contradiction that there exists an $\tilde{\alpha}$ such that

$$F(\tilde{\alpha}, z^*(\tilde{\alpha})) = G(\tilde{\alpha}, z^*(\tilde{\alpha})) > \inf_z \{\max\{F(\tilde{\alpha}, z), G(\tilde{\alpha}, z)\}\}. \tag{4.4.6}$$

Then there exists $\epsilon > 0$ such that

$$F(\tilde{\alpha}, z^*(\tilde{\alpha})) = G(\tilde{\alpha}, z^*(\tilde{\alpha})) \geq \inf_z \{\max\{F(\tilde{\alpha}, z), G(\tilde{\alpha}, z)\}\} + \epsilon,$$

which implies that there exists $\hat{z}(\tilde{\alpha}, \epsilon) \in \mathbb{R}$ such that

$$F(\tilde{\alpha}, z^*(\tilde{\alpha})) = G(\tilde{\alpha}, z^*(\tilde{\alpha})) \geq \max\{F(\tilde{\alpha}, \hat{z}(\tilde{\alpha}, \epsilon)), G(\tilde{\alpha}, \hat{z}(\tilde{\alpha}, \epsilon))\}.$$

Since F is non-increasing in z the above inequality implies

$$F(\tilde{\alpha}, z^*(\tilde{\alpha})) \geq F(\tilde{\alpha}, \hat{z}(\tilde{\alpha}, \epsilon)) \Rightarrow \hat{z}(\tilde{\alpha}, \epsilon) \geq z^*(\tilde{\alpha}),$$

and similarly since G is non-decreasing in z

$$G(\tilde{\alpha}, z^*(\tilde{\alpha})) \geq G(\tilde{\alpha}, \hat{z}) \Rightarrow z^*(\tilde{\alpha}) \geq \hat{z}(\tilde{\alpha}, \epsilon).$$

Hence $z^*(\tilde{\alpha}) = \hat{z}(\tilde{\alpha}, \epsilon)$ yields a contradiction to (4.4.6).

Remark 4.4.2. Let $\mathcal{A} \subseteq \mathbb{R}$ and $F, G : \mathcal{A} \times \mathbb{R} \to \mathbb{R}$ satisfy the properties in (4.4.1). Assume that for any α, there exists at least one $z^*(\alpha)$ such that $F(\alpha, z^*(\alpha)) = G(\alpha, z^*(\alpha))$ then the following holds

$$\inf_{\alpha,z}\{\max\{F(\alpha,z),G(\alpha,z)\}\} = \inf_{\alpha} F(\alpha,z^*(\alpha)) = \inf_{\alpha} G(\alpha,z^*(\alpha)). \tag{4.4.7}$$

Theorem 4.4.3. $v(t,x,u)$ is convex in u for all x and fixed t.

Proof. By definition of v we know that $\forall\, \epsilon > 0 \quad \exists\ y^{i,\epsilon}, Z^{i,\epsilon}, \alpha^{i,\epsilon}$ for $i = 1,2$ such that

$$\begin{aligned} y^{i,\epsilon} &< v(t,x,\alpha^i) + \epsilon, \\ Y_T^{t,y^{i,\epsilon},Z^{i,\epsilon}} &\geq U^{-1}(u_t^{t,u^i,\alpha^{i,\epsilon}}) + g(X_T^{t,x}). \end{aligned}$$

Set

$$\begin{aligned} Z^\epsilon &= \lambda Z^{1,\epsilon} + (1-\lambda)Z^{2,\epsilon}, \\ \alpha^\epsilon &= \lambda\alpha^{1,\epsilon} + (1-\lambda)\alpha^{2,\epsilon}, \\ y^\epsilon &= \lambda y^{1,\epsilon} + (1-\lambda)y^{2,\epsilon}. \end{aligned}$$

We claim that $\quad Y_T^{t,y^\epsilon,Z^\epsilon} \geq U^{-1}\left(u_T^{t,\lambda u_1+(1-\lambda)u_2,\alpha^\epsilon}\right) + g(X_T^{t,x}).$
Indeed,

$$\begin{aligned} Y_T^{t,y^\epsilon,Z^\epsilon} &= y^\epsilon + \sum_{j=t}^{T-1} Z_j^\epsilon(X_{j+1}-X_j), \\ &= \lambda y^{1,\epsilon} + (1-\lambda)y^{2,\epsilon} + \sum_{j=t}^{T-1}\left(\lambda Z_j^{1,\epsilon} + (1-\lambda)Z_j^{2,\epsilon}\right)(X_{j+1}-X_j), \\ &= \lambda\left(Y_T^{y_1^{t,\epsilon},Z_1^\epsilon}\right) + (1-\lambda)\left(Y_T^{t,y_2^\epsilon,Z_2^\epsilon}\right), \\ &\geq \lambda\left(U^{-1}(u_T^{t,u_1,\alpha_1^\epsilon}) + g(X_T^{t,x})\right) + (1-\lambda)\left(U^{-1}(u_T^{t,u_2,\alpha_2^\epsilon}) + g(X_T^{t,x})\right), \\ &\geq U^{-1}\left(u_T^{t,\lambda u_1+(1-\lambda)u_2,\alpha^\epsilon}\right) + g(X_T^{t,x}) \qquad \text{(since } U^{-1} \text{ is convex in u)}\ . \end{aligned}$$

Clearly, $\lambda y^{1,\epsilon} < \lambda v(t,x,u^1) + \lambda\epsilon$ and $(1-\lambda)y^{2,\epsilon} < (1-\lambda)v(t,x,u^2) + (1-\lambda)\epsilon$.

Thus, $y^{\epsilon} = \lambda y^{1,\epsilon} + (1-\lambda)y^{2,\epsilon} \leq \lambda v(t,x,u^1) + (1-\lambda)v(t,x,u^2) + \epsilon$.
Let $\epsilon \to 0$, then $v\left(t,x,\lambda u_1 + (1-\lambda)u_2\right) \leq \lambda v(t,x,u_1) + (1-\lambda)v(t,x,u_2)$.

After proving the convexity of the value function, now let us turn back to our dynamic programming equation (4.4.4). If we apply Lemma (4.4.1) together with

$$\begin{aligned} F(\alpha,z) &= v\left(t+1,x_{up},u+\alpha\right) - z(x_{up}-x), \\ G(\alpha,z) &= v\left(t+1,x_{down},u-\alpha\right) - z(x_{down}-x), \end{aligned}$$

and setting $F(\alpha,z) = G(\alpha,z)$

$$\Rightarrow z^*(\alpha) = \frac{v\left(t+1,x_{up},u+\alpha\right) - v\left(t+1,x_{down},u-\alpha\right)}{x_{up}-x_{down}}.$$

Hence,

$$\begin{aligned} v(t,x,u) &= F(\alpha,z^*(\alpha)) = G(\alpha,z^*(\alpha)) \\ &= \inf_{\alpha}\left(\frac{(x-x_{down})v\left(t+1,x_{up},u+\alpha\right) + (x_{up}-x)v\left(t+1,x_{down},u-\alpha\right)}{x_{up}-x_{down}}\right) \end{aligned}$$

Let $p^* = \left(\frac{x-x_{down}}{x_{up}-x_{down}}\right)$. Then, $v(t,x,u)$ can also be written as

$$v(t,x,u) = \inf_{\alpha}\{p^* v\left(t+1,x_{up},u+\alpha\right) + (1-p^*)v\left(t+1,x_{down},u-\alpha\right)\} \tag{4.4.8}$$

By using the no-arbitrage condition (4.2.2), we obtain a sum of two convex functions in (4.4.8) which is another convex function. This helps considerably to construct an efficient algorithm for the solution.

4.5 Application to Quantile Hedging

4.5.1 The formulation of the quantile hedging problem for complete markets

The idea of quantile hedging is formulated in [10] as the following. It is assumed that the discounted price process of the underlying is a semi-martingale $X = (X_t)_{t\in[0,T]}$ on a probability space $(\Omega, \mathcal{F}, P)$ with the filtration $(\mathcal{F}_t)$. P denotes the set of all equivalent martingale measures and absence of arbitrage implies $P \neq \emptyset$. In a complete market, there exists an unique equivalent martingale measure $P^* \approx P$. Now, consider a contingent claim given by an $\mathcal{F}_T$-measurable, nonnegative random variable H such that $H \in L^1(P^*)$. Completeness implies the existence of a predictable process ξ^H, providing a perfect hedge for H i.e.

$$E^*[H|F_t] = H_0 + \int_0^t \xi_s^H dX_s \quad \forall t \in [0,T] \quad P - a.s$$

where E^* denotes the expectation with respect to P^*. Thus, the claim can be replicated by the self-financing trading strategy (H_0, ξ^H). This already assumes that the investor is ready to allocate the required initial capital $H_0 = E^*[H]$. However, the investor may be unwilling or unable to invest the initial capital H_0. In this case, the solution to the quantile hedging problem provides the best hedge the investor can achieve with a given smaller amount $\widetilde{V_0} < H_0$. The probability that the hedge ie successful is taken as the optimality criterion. More precisely, the solution constructs an admissible strategy (V_0, ξ^*) such that the corresponding value process V^* satisfies

$$P\left[V_T^* = V_0 + \int_0^T \xi_s^* dX_s \geq H\right] = \max P\left[V_T = V_0 + \int_0^T \xi_s dX_s \geq H\right],$$

where the maximum is over the set of all admissible portfolio strategies satisfying

$$V_0 \leq \widetilde{V_0}.$$

The probability $P\left[V_T \geq H\right]$ is termed as the success probability.

4.5.2 Solution by our method

In the original problem

$$V(t,x) \quad := \quad \inf\left\{\ y\ :\ \exists\{\mathrm{Z}\} \in \mathcal{A} \quad \text{s.t.} \quad \mathrm{E}\left[U\left(Y_T^{t,y,Z} - g(X_T^{t,x})\right)\right] \geq u^*\ \right\},$$

we take

$$U(x) = \begin{cases} 1, & \text{if } x \geq 0, \\ -\infty, & \text{if } x < 0, \end{cases}$$

to obtain

$$V(t,x) \quad := \quad \inf\left\{\ y\ :\ \exists\{\mathrm{Z}\} \in \mathcal{A} \quad \text{s.t.} \quad P\left(Y_T^{t,y,Z} \geq g(X_T^{t,x})\right) \geq u^*\ \right\}.$$

where $u^* := p^*$ is the success probability.

4.5.3 Quantile Hedging in a Market with Frictions

A general definition of financial market frictions is provided by DeGennaro and Robotti [2] as anything that interferes with trades that rational individuals make (or would make in the absence of market frictions). The sources of financial market frictions are diverse and widespread but still can be classified into (although not completely exhaustive) transactions costs, portfolio constraints, taxes and regulations, differential borrowing and lending rates, asset indivisibility, and agency and information problems. In this section we focus on the portfolio constraints type of financial market frictions. The effects of leverage (or portfolio) constraints on optimal hedging of stock and bond options in discrete time is first studied in Naik and Uppal [13]. Then Broadie et al. [14] extend the work of [13] to the continuous-time framework. We use their results as a benchmark for our model. As a result, we can model the borrowing constraint by requiring that the amount borrowed cannot exceed C_b times total wealth, i.e.,

$$(Z_t X_t - Y_t) \leq C_b Y_t \Rightarrow Z_t \leq \frac{(1+C_b)Y_t}{X_t}, \qquad 0 \leq t \leq T. \tag{4.5.1}$$

Similarly, we require short selling amount to be less than C_s times total wealth by imposing the short selling constraint

$$-Z_t X_t \leq C_s X_t \Rightarrow Z_t \geq \frac{-C_s Y_t}{X_t}, \qquad 0 \leq t \leq T. \tag{4.5.2}$$

4.5.4 Dynamic Programming with Portfolio Constraints

Under the portfolio process restrictions (4.5.1) and (4.5.2), the dynamic programming equation in (4.4.4) can be stated as

$$v(t,x,u) = \inf_{z,\alpha}\{max\{F(\alpha,z), G(\alpha,z)\}\},$$

where z takes values in the set $\left[\frac{-C_s v(t,x,u)}{x}, \frac{(1+C_b)v(t,x,u)}{x}\right]$ and

$$\begin{aligned} F(\alpha,z) &:= v\left(t+1, x_{up}, u+\alpha\right) - z(x_{up}-x), \\ G(\alpha,z) &:= v\left(t+1, x_{down}, u-\alpha\right) - z(x_{down}-x). \end{aligned}$$

Note that the appearance of the value function in the boundaries of z gives the optimization problem an implicit nature. In order to solve this problem, we first fix α which is assumed to be the minimizer. Then we calculate $z^*(\alpha)$ according to Lemma(4.4.1) as

$$z^*(\alpha) = \frac{v\left(t+1, x_{up}, u+\alpha\right) - v\left(t+1, x_{down}, u-\alpha\right)}{2x\sigma\sqrt{h}}.$$

We differentiate the following cases:

Case 1: $\quad \frac{-C_s v(t,x,u)}{x} < z^*(\alpha) < \frac{(1+C_b)v(t,x,u)}{x}$.

In this case, the problem is the same as the original problem which does not have any portfolio constraints. Hence, we set $z^{c,*}(\alpha) = z^*(\alpha)$ where $z^{c,*}(\alpha)$ is the optimal portfolio in case of the portfolio constraints. Then we define a new function H such that

$$H(\alpha) = F(\alpha, z^{c,*}) = G(\alpha, z^{c,*}), \tag{4.5.3}$$

yielding

$$H(\alpha) = \frac{v\left(t+1, x_{up}, u+\alpha\right)\left(x_{down}-x\right) + v\left(t+1, x_{down}, u-\alpha\right)\left(x_{up}-x\right)}{2x\sigma\sqrt{h}}.$$

Case 2: $\quad z^*(\alpha) \geq \frac{(1+C_b)v(t,x,u)}{x}$.

In this case, we set the optimal stock amount to the upper boundary of the interval

$$z^{c,*}(\alpha) = \frac{(1+C_b)v(t,x,u)}{x},$$

which implies that

$$\begin{aligned} v(t,x,u) &= \max\left\{F(\alpha, z^{c,*}), G(\alpha, z^{c,*})\right\}, \\ &= \max\left\{\begin{array}{l} v\left(t+1, x_{up}, u+\alpha\right) - (1+C_b)v(t,x,u)(\mu h + \sigma\sqrt{h}), \\ v\left(t+1, x_{down}, u-\alpha\right) - (1+C_b)v(t,x,u)(\mu h - \sigma\sqrt{h}). \end{array}\right\} \end{aligned}$$

If $F(\alpha, z^{c,*})$ is the maximum then

$$v(t,x,u) = v\left(t+1, x_{up}, u+\alpha\right) - (1+C_b)v(t,x,u)(\mu h + \sigma\sqrt{h}),$$

$$\Rightarrow \quad v(t,x,u) = \frac{v\left(t+1, x_{up}, u+\alpha\right)}{1 + (1+C_b)(\mu h + \sigma\sqrt{h})}.$$

If $G(\alpha, z^{c,*})$ is the maximum then

$$v(t,x,u) = v\left(t+1, x_{down}, u-\alpha\right) - (1+C_b)v(t,x,u)(\mu h - \sigma\sqrt{h}),$$

$$\Rightarrow \quad v(t,x,u) = \frac{v\left(t+1, x_{down}, u-\alpha\right)}{1+(1+C_b)(\mu h - \sigma\sqrt{h})}.$$

Then we set

$$H(\alpha) = \max\left\{\frac{v\left(t+1, x_{up}, u+\alpha\right)}{1+(1+C_b)(\mu h + \sigma\sqrt{h})}, \frac{v\left(t+1, x_{down}, u-\alpha\right)}{1+(1+C_b)(\mu h - \sigma\sqrt{h})}\right\}.$$

Claim: $H(\alpha) = \frac{v(t+1, x_{up}, u+\alpha)}{1+(1+C_b)(\mu h + \sigma\sqrt{h})}$ in the above maximization problem.

Proof.

By using the constraint in Case 2,

$$\begin{aligned}
z^*(\alpha) &\geq \frac{(1+C_b)v(t,x,u)}{x}, \\
&\Rightarrow \frac{v\left(t+1, x_{up}, u+\alpha\right) - v\left(t+1, x_{down}, u-\alpha\right)}{2x\sigma\sqrt{h}}, \\
&\geq \frac{(1+C_b)}{x}\left(\frac{v\left(t+1, x_{up}, u+\alpha\right)\left(\sigma\sqrt{h} - \mu h\right)}{2\sigma\sqrt{h}} + \right. \\
&\qquad \left. \frac{v\left(t+1, x_{down}, u-\alpha\right)\left(\sigma\sqrt{h} + \mu h\right)}{2\sigma\sqrt{h}}\right), \\
&\Rightarrow \left(1+(1+C_b)(\mu h - \sigma\sqrt{h})\right) v\left(t+1, x_{up}, u+\alpha\right), \\
&\geq \left(1+(1+C_b)(\mu h + \sigma\sqrt{h})\right) v\left(t+1, x_{down}, u-\alpha\right), \\
&\Rightarrow \frac{v\left(t+1, x_{up}, u+\alpha\right)}{\left(1+(1+C_b)(\mu h + \sigma\sqrt{h})\right)} c \geq \frac{v\left(t+1, x_{down}, u-\alpha\right)}{\left(1+(1+C_b)(\mu h - \sigma\sqrt{h})\right)}.
\end{aligned}$$

Case 3: $z^*(\alpha) \leq \frac{-C_s v(t,x,u)}{x}$

In this case, we set $z^{c,*}(\alpha) = \frac{-C_s v(t,x,u)}{x}$. Then

$$\begin{aligned} v(t,x,u) &= \max\left\{F(\alpha, z^{c,*}), G(\alpha, z^{c,*})\right\}, \\ &= \max\left\{\begin{array}{l} v\left(t+1, x_{up}, u+\alpha\right) + C_s v(t,x,u)(\mu h + \sigma\sqrt{h}), \\ v\left(t+1, x_{down}, u-\alpha\right) + C_s v(t,x,u)(\mu h - \sigma\sqrt{h}) \end{array}\right\}. \end{aligned}$$

If $F(\alpha, z^{c,*})$ is the maximum then

$$\begin{aligned} v(t,x,u) &= v\left(t+1, x_{up}, u+\alpha\right) + C_s v(t,x,u)(\mu h + \sigma\sqrt{h}), \\ &\Rightarrow \quad v(t,x,u) = \frac{v\left(t+1, x_{up}, u+\alpha\right)}{1 - C_s(\mu h + \sigma\sqrt{h})}. \end{aligned}$$

If $G(\alpha, z^{c,*})$ is the maximum then

$$\begin{aligned} v(t,x,u) &= v\left(t+1, x_{down}, u-\alpha\right) + C_s v(t,x,u)(\mu h - \sigma\sqrt{h}), \\ &\Rightarrow \quad v(t,x,u) = \frac{v\left(t+1, x_{down}, u-\alpha\right)}{1 - C_s(\mu h - \sigma\sqrt{h})}. \end{aligned}$$

Hence, we set

$$H(\alpha) = \max\left\{\frac{v\left(t+1, x_{up}, u+\alpha\right)}{1 - C_s(\mu h + \sigma\sqrt{h})}, \frac{v\left(t+1, x_{down}, u-\alpha\right)}{1 - C_s(\mu h - \sigma\sqrt{h})}\right\}.$$

Claim: $H(\alpha) = \frac{v(t+1, x_{down}, u-\alpha)}{1-C_s(\mu h - \sigma\sqrt{h})}$ in the above maximization problem.

Proof.

From the constraint in Case 3,

$$
\begin{aligned}
z^*(\alpha) &\le \frac{-C_s v(t,x,u)}{x}, \\
&\Rightarrow \frac{v\left(t+1,x_{up},u+\alpha\right) - v\left(t+1,x_{down},u-\alpha\right)}{2x\sigma\sqrt{h}}, \\
&\le \frac{(-C_s)}{x}\left(\frac{v\left(t+1,x_{up},u+\alpha\right)(\sigma\sqrt{h}-\mu h)}{2\sigma\sqrt{h}} + \right. \\
&\qquad \left. \frac{v\left(t+1,x_{down},u-\alpha\right)(\sigma\sqrt{h}+\mu h)}{2\sigma\sqrt{h}}\right), \\
&\Rightarrow \left(1 - C_s(\mu h + \sigma\sqrt{h})\right) v\left(t+1,x_{down},u-\alpha\right), \\
&\ge \left(1 + C_s(\sigma\sqrt{h} - \mu h)\right) v\left(t+1,x_{up},u+\alpha\right), \\
&\Rightarrow \frac{v\left(t+1,x_{down},u-\alpha\right)}{\left(1 - C_s(\mu h - \sigma\sqrt{h})\right)} \ge \frac{v\left(t+1,x_{up},u+\alpha\right)}{\left(1 - C_s(\mu h + \sigma\sqrt{h})\right)}.
\end{aligned}
$$

We follow this case by case analysis for each α and then we take minimum over α and set the value function to this value

$$v(t,x,u) = \min_{\alpha} H(\alpha).$$

4.5.5 Numerical Results

In this section, we present the numerical results of our model for European Vanilla and barrier call and put options. Moreover, we investigate the quantile hedging costs under portfolio constraints.

As a benchmark, we utilize the closed form solutions for Vanilla call and put options derived in [10] to test our method. They provide analytical solution for the case of a call option when $\mu < \sigma^2$. Hence, in order to use the analytical solution we take volatility: $\sigma = 0.3$; mean return rate: $\mu = 0.08$ as in [10] for European call options. For the other parameters we use the initial stock price: $S_0 = 100$, strikes $K = 90, 100, 110$; maturities: $T = 1$ month, 3-month, 6-month, 9-month, 12-month; number of time steps: $N = 100$, time discretization $h = 0.001$, and for the simplicity of computations interest rate is taken to be zero. We consider the shortfall probability up to 10 % as a reasonable level of probability to accept and we conduct our numerical experiments according to that. Figures (4.1), (4.2), and (4.3) show the absolute percentage errors for in-the-money (ITM), at-the-money (ATM), out-of-the-money (OTM) call options computed with our method. As is clearly visible from the figures, as the shortfall probability changes from 0 to 10 %, the absolute percentage error always remains less than 1 % for almost all of the parameter values. The only exception is the 1-month OTM option which is not shown in the figure. However as the number of time steps increases it is possible to produce more accurate results with very small percentage errors. One can obtain similar results for put options.

The fourth column in Tables (4.1)-(4.3), shows how much percentage gain an investor can make by accepting a certain shortfall probability. For instance, for a 6-month ATM call option it is possible to pay 5.5 % less by accepting a 1 % shortfall probability and this increases up to 39.5 % if one accepts a 10 % shortfall probability. We observe that percentage gains increase with time to maturity for ITM call options and decrease for ATM and OTM call options. There is also numerical evidence in Tables (4.1)-(4.3) that an inverse relationship exists between the moneyness of the call option and profit from the quantile hedging strategy. The above relations hold in the reverse direction for put options. (Numerical results for put options are available upon request.)

If strict borrowing constraints are imposed on the wealth process, then percentage gains from the quantile hedging approach become substantial for the short term ITM, ATM, OTM call options. For example, when an investor is allowed to borrow at most two times more than his current wealth i.e., when $C_b = 2$, then if he accepts a 1 % shortfall risk for an ATM call option, he can pay 23.7 % less than the BS-price. This is almost four times more profitable than in the unconstrained case. However, as the shortfall probability increases the effect of portfolio constraints on the relative profitability of the quatile hedging strategy diminishes. If we take a 10 % shortfall size in the previous example, then the constrained case is only at most 1.5 times more profitable than the unconstrained case. One can also observe from the Figures (4.4)-(4.6) that as time to maturity increases the effect of borrowing constraints becomes less and less significant for all options independent of the moneyness of the option.

Relaxing the portfolio constraints on the wealth process, i.e. increasing C_b then the numbers converge to the numbers in the unconstrained case as expected. As a test case, we take $C_b = 1000$ which yields exactly the same numbers as the unconstrained case. For a second benchmark, we use the explicit formulas given in Broadie et al. [14] for the minimum super-replication prices under portfolio constraints in continuous time. Their methodology is to first create a dominating claim which has appropriately increased payoffs with respect to the original claim. Then they prove that the price of the original claim with portfolio constraints is the price of the dominating claim without constraints. In particular, the dominating claim for a standard call option with payoff function $b(X) = (X-K)^+$ is given by

$$\hat{b}(X) = \begin{cases} X - K, & \text{if } X \geq \frac{Ku}{u-1}, \\ \frac{K}{u-1}\left(\frac{(u-1)X}{Ku}\right)^u, & \text{if } X < \frac{Ku}{u-1}, \end{cases} \tag{4.5.4}$$

where u is the borrowing constraint which corresponds to $(1 + C_b)$ in our context. Let $\hat{y}$ denote the minimum super-replication call price with borrowing constraints calculated from the formulas in (4.5.4) on a binomial tree with 100 time steps. Let $\tilde{y}$ denote the minimal initial capital required under borrowing constraints when the shortfall probability is zero. We compute $\tilde{y}$ with our method again in 100 time steps. A list of $\hat{y}$ and $\tilde{y}$ values are presented in Table (4.4) for different call options. A comparison of $\tilde{y}$ and $\hat{y}$ values in Table (4.4) demonstrates that our method provides a very close approximation to the continuous time model. The difference stems mainly from the discretization procedure.

Quantile hedging costs for exotic options can also be computed with our method. As an example, we consider an up-and-out barrier call option with the parameters given in Table (4.5). Options with barriers are constructed to decrease the initial cost of a similar option without barrier. Quantile hedging strategies can further decrease the initial cost more. The results in Table (4.5) shows that the percentage gains increase with time to maturity independent of the moneyness of the option. We also observe an inverse relationship between the moneyness of the call option and the profit from the quantile hedging strategy.

4.6 Other applications

Our method provides a unified approach to seemingly different finance problems such as utility indifference pricing, good deal pricing, and the expected shortfall problem.

4.6.1 Application to Utility Indifference Pricing

Utility based pricing is a theoretically appealing pricing methodology in incomplete markets. It has been introduced by Hodges and Neuberger [3] and extensively studied in the recent finance literature. As it is clearly described in Musiela and Zariphopoulou [6] the indifference pricing provides a link between pricing a derivative product and maximization of utility of wealth which is important not only for sell side (such as investment banks) but also for the buy side (such as wealth managers) of the financial markets. Inclusion of risk aversion and wealth dependence in utility indifference pricing makes it an economically natural and justified method in incomplete markets. Carassus and Rásonyi [8] define the seller's indifference price as the minimal amount a seller should add to her initial wealth so as to reach an optimal expected utility when delivering the claim which is greater than or equal to the one she would have obtained trading in the basic assets only.

Let p^s denote the seller's indifference price. Considering the problem with λ units of claim, it can be formalized in our setting as the following

$$\max_{Z\in\mathcal{A}^t} \mathrm{E}\left[U\left(Y_N^{t,y+\lambda p^s,Z} - \lambda g(X_T^{t,x})\right)\right] = \max_{Z\in\mathcal{A}^t} \mathrm{E}\left[U\left(Y_T^{t,y,Z}\right)\right] \tag{4.6.1}$$

Similarly buyer's price p^b can be obtained from

$$\max_{Z\in\mathcal{A}^t} \mathrm{E}\left[U\left(Y_N^{t,y-\lambda p^b,Z} + \lambda g(X_T^{t,x})\right)\right] = \max_{Z\in\mathcal{A}^t} \mathrm{E}\left[U\left(Y_T^{t,y,Z}\right)\right] \tag{4.6.2}$$

In the original problem (4.3.1), we start with a strictly increasing and concave utility function U but the threshold u^* is also determined by this utility function

$$u^*(t,y) = \sup_{Z\in\mathcal{A}^t} \mathrm{E}\left[U\left(Y_T^{t,y,Z}\right)\right]. \tag{4.6.3}$$

4.6.2 Application to Good-Deal Bounds

Cochrane and SaáRequejo [5] describe the concept of a *good deal* as an investment opportunity with a *too* high Sharpe ratio. Their idea is to rule out good deals by putting a bound on the Sharpe ratios. This bound is called a *good deal bound.* The simultaneously published paper by Bernardo and Ledoit [23] use *gain-loss ratio* instead of Sharpe ratio. The expectation of the investment's positive excess payoffs divided by the expectation of its negative excess payoffs, $\frac{E[y^+]}{E[y^-]}$, is defined as the *gain loss ratio* to measure the attractiveness of an investment opportunity. They impose a bound, λ, on the gain loss ratio to obtain tighter price bounds than no-arbitrage price bounds. Similar to [23]'s gain loss ratio, Pınar et al. [24] develop *sufficiently attractive expected gain* (SAGE) approach. A sequence of portfolio holdings is said to yield a SAGE at level $\lambda > 1$ if

$$E[y^+] - \lambda E[y^-] > 0,$$

where λ is the loss aversion parameter of an individual investor. In our context, we define

$$U(y) = y^+ - \lambda y^-,$$

where y^+ and y^- are the positive and negative parts of y and we take $u^* = 0$ in the general problem (4.3.1) then we formulate the seller's price, S^λ, as

$$\mathbf{S}^\lambda \quad := \quad \inf\left\{ y \ : \ \exists\{Z\} \in \mathcal{A}^t \text{ s.t. } \mathrm{E}\left[\left(Y_T^{t,y,Z} - g(X_T^{t,x})\right)^+ - \lambda\left(Y_T^{t,y,Z} - g(X_T^{t,x})\right)^-\right] \geq 0 \right\}$$

The above definition can also be written in terms of the equivalent formulation of the original problem as

$$\mathbf{S}^\lambda \quad := \quad \inf\left\{ y \ : \ \exists\{Z\} \in \mathcal{A}^t, \ \{\alpha\}, \text{ s.t.} \quad Y_T^{t,y,Z} \geq U^{-1}(u_T^{t,u^*,\alpha}) + g(X_T^{t,x}) \ a.s. \right\},$$

where

$$U^{-1}(u) = \begin{cases} u, & \text{if } u \geq 0, \\ \frac{u}{\lambda}, & \text{if } u < 0, \end{cases}$$

with the terminal condition in the dynamic programming

$$v(T,x,u) = \begin{cases} u + g(x), & \text{if } u \geq 0, \\ \frac{u}{\lambda} + g(x), & \text{if } u < 0. \end{cases}$$

Similarly, the buyer's problem can be formulated as

$$\mathbf{B}^{\lambda} \quad := \quad \sup\left\{ y \ : \ \exists\{\tilde{Z}\} \in \mathcal{A}^t \text{ s.t. E} \left[\left(-Y_N^{t,y,\tilde{Z}} + g(X_T^{t,x})\right)^+ - \lambda \left(-Y_N^{t,y,\tilde{Z}} + g(X_T^{t,x})\right)^- \right] \geq 0 \right\}.$$

4.6.3 Application to Expected Shortfall

Value at risk (VaR) is defined as the worst expected loss over a given time period at a given confidence level under normal market conditions. VaR is a popular risk measure for fund managers, corporate treasures, banks and other financial institutions. Its wide use can be attributed to mandatory regulations endorsed by the Basel Committee on Banking Supervision. However, as it is clear from its definition VaR considers only probability of loss but not the size of it. Artzner et al. [25] introduce an alternative risk measure Expected Shortfall (ES) which takes both the size and probability of shortfall into account. They propose four axioms which should be satisfied by a coherent risk measure and show that ES satisfies three of four axioms. In contrast to ES, VaR does not satisfy sub-additivity axiom which is an important feature for a reasonable risk measure. A violation of this axiom means that portfolio diversification may result in an increase of risk which contradicts with the general risk concept. In our context, we fix a bound on the expected shortfall size

$$E\left[max\left(0, g(X_T^{t,x}) - Y_T^{t,y,Z}\right)\right],$$

and try to minimize required initial amount. In the general problem (4.3.1), we take $U(x) = -max(0,-x)$ then

$$V(t,x) := \inf\left\{ y \ : \ \exists\{Z\} \in \mathcal{A}^t \text{ s.t. } \ E\left[max\left(0, g(X_T^{t,x}) - Y_T^{t,y,Z}\right)\right] \leq -u^* \right\},$$

In terms of the equivalent formulation of the original problem

$$v(t,x,u) \;:=\; \inf\Big\{\, y \;:\; \exists\{Z\}\in\mathcal{A}^t,\ \{\alpha\},\ \text{s.t.}\\ Y_T^{t,y,Z} \geq U^{-1}(u_T^{t,u^*,\alpha}) + g(X_T^{t,x})\ a.s.\,\Big\},$$

where

$$U^{-1}(u) = \begin{cases} \infty, & \text{if } u \geq 0,\\ u, & \text{if } u < 0,\end{cases}$$

with the terminal condition in the dynamic programming

$$V(T,x,u) = \begin{cases} \infty, & \text{if } u \geq 0,\\ u + g(x), & \text{if } u < 0.\end{cases}$$

4.7 Conclusion

In this paper, we consider a discrete-time version of the stochastic target problem which developed by [11] in continuous-time. Using the standard techniques in [20] and [15], we develop the dynamic programming equation for the general problem and the problem with portfolio constraints. Our discrete-time model provides a unified alternative approach for the solutions of different finance problems such as quantile hedging, utility indifference pricing, good deal pricing, and expected shortfall. However, since the other problems have been extensively studied in the literature, we present an application of our method to the quantile hedging problem. Our efficient algorithm provides a detailed numerical analysis for the quantile hedging of Vanilla and exotic options in addition to the problem with portfolio constraints.

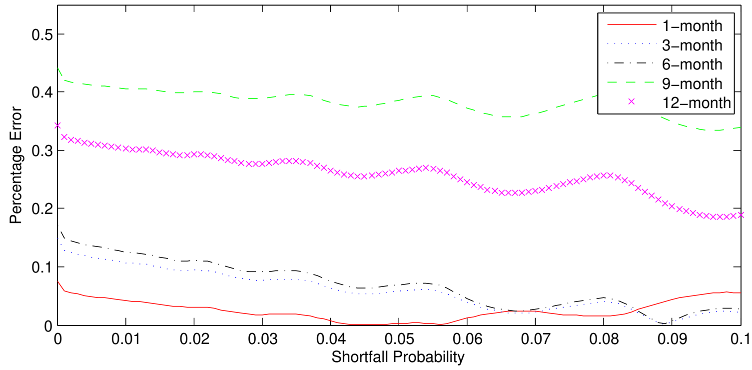

Figure 4.1: Call Option with $S = 100$, $K = 90$, $\sigma = 0.3$, $\mu = 0.08$

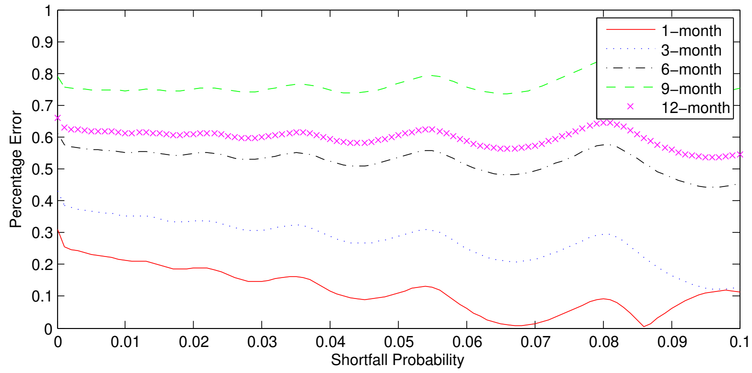

Figure 4.2: Call Option with $S = 100$, $K = 100$, $\sigma = 0.3$, $\mu = 0.08$

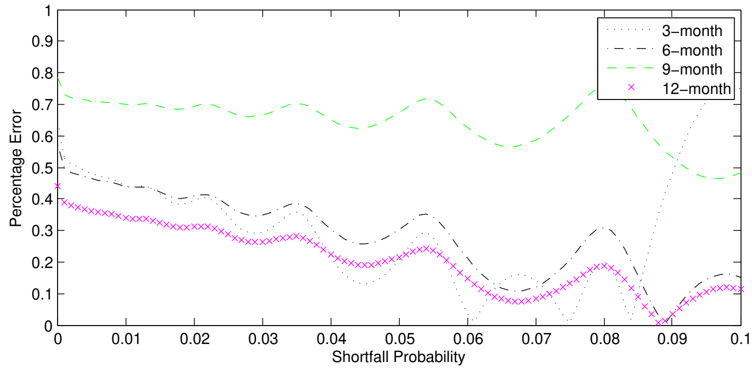

Figure 4.3: Call Option with $S = 100$, $K = 110$, $\sigma = 0.3$, $\mu = 0.08$

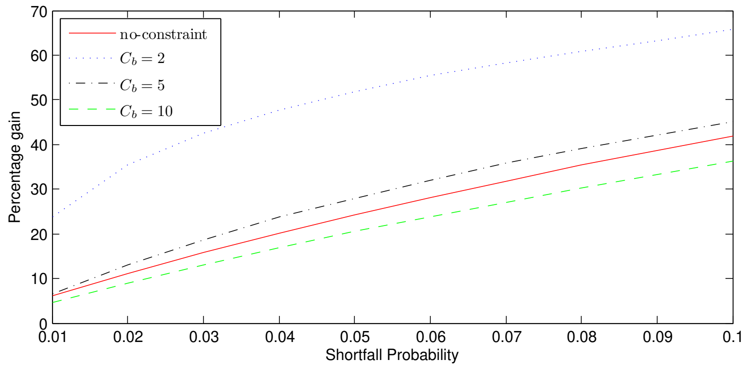

Figure 4.4: Call Option with $S = 100$, $K = 90$, $\sigma = 0.3$, $\mu = 0.08$

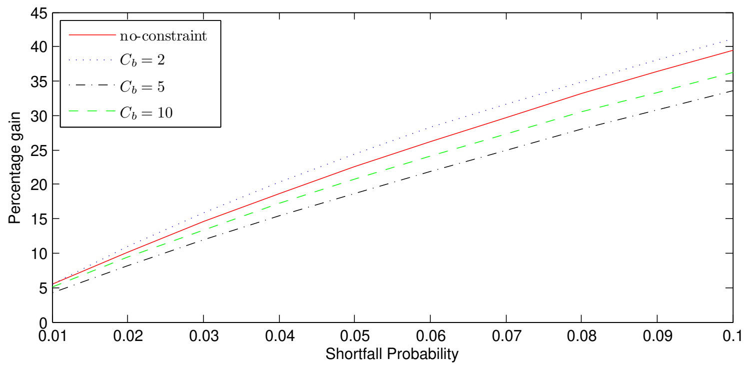

Figure 4.5: Call Option with $S = 100$, $K = 100$, $\sigma = 0.3$, $\mu = 0.08$

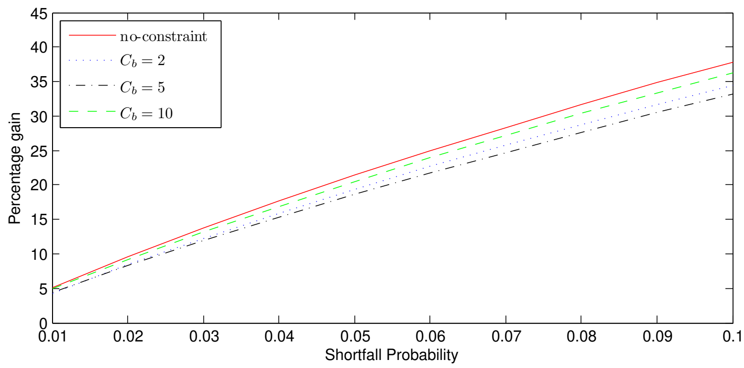

Figure 4.6: Call Option with $S = 100$, $K = 110$, $\sigma = 0.3$, $\mu = 0.08$

Table 4.1: Quantile Hedging costs for a call option under borrowing constraint with parameters $S_0 = 100$, $K = 90$, $\mu = 0.08$, $\sigma = 0.3$.

T	Shortfall Prob	Initial Cost	% Gain	C_b 2	% Gain	C_b 5	% Gain	C_b 10	% Gain	C_b 1000
0.083	0	10.44		18.70		12.45		11.02		10.44
0.083	0.01	10.15	2.8	16.89	9.7	12.13	2.6	10.73	2.6	10.15
0.083	0.02	9.89	5.3	15.65	16.3	11.82	5.1	10.47	5.0	9.89
0.083	0.03	9.64	7.6	14.78	21.0	11.53	7.4	10.23	7.2	9.64
0.083	0.04	9.40	9.9	14.10	24.6	11.25	9.7	9.99	9.4	9.40
0.083	0.05	9.17	12.1	13.53	27.7	10.98	11.8	9.76	11.5	9.17
0.083	0.06	8.95	14.2	13.03	30.3	10.72	13.9	9.54	13.5	8.95
0.083	0.07	8.74	16.3	12.59	32.7	10.47	15.9	9.32	15.4	8.74
0.083	0.08	8.53	18.3	12.18	34.9	10.23	17.8	9.11	17.3	8.53
0.083	0.09	8.32	20.3	11.79	37.0	9.99	19.7	8.91	19.2	8.32
0.083	0.10	8.12	22.2	11.44	38.8	9.76	21.6	8.71	21.0	8.12
0.5	0	13.97		20.72		15.76		14.52		13.97
0.5	0.01	13.44	3.7	19.99	3.5	15.24	3.3	14.00	3.6	13.44
0.5	0.02	12.98	7.1	19.26	7.0	14.77	6.3	13.54	6.8	12.98
0.5	0.03	12.55	10.2	18.58	10.4	14.34	9.0	13.10	9.8	12.55
0.5	0.04	12.13	13.1	17.92	13.5	13.92	11.7	12.69	12.6	12.13
0.5	0.05	11.74	16.0	17.32	16.4	13.52	14.2	12.29	15.3	11.74
0.5	0.06	11.36	18.7	16.75	19.1	13.13	16.7	11.91	18.0	11.36
0.5	0.07	10.99	21.3	16.20	21.8	12.76	19.0	11.54	20.5	10.99
0.5	0.08	10.62	23.9	15.67	24.4	12.39	21.4	11.18	23.0	10.62
0.5	0.09	10.28	26.4	15.19	26.7	12.03	23.7	10.84	25.4	10.28
0.5	0.10	9.94	28.8	14.71	29.0	11.69	25.8	10.50	27.7	9.94
1	0	16.95		22.80		18.46		17.40		16.95
1	0.01	16.30	3.9	22.04	3.4	17.81	3.5	16.74	3.8	16.30
1	0.02	15.71	7.3	21.34	6.4	17.22	6.7	16.16	7.1	15.71
1	0.03	15.17	10.6	20.66	9.4	16.67	9.7	15.61	10.3	15.17
1	0.04	14.64	13.6	20.01	12.2	16.15	12.5	15.09	13.3	14.64
1	0.05	14.14	16.6	19.41	14.9	15.65	15.2	14.58	16.2	14.14
1	0.06	13.66	19.5	18.82	17.5	15.16	17.9	14.10	19.0	13.66
1	0.07	13.19	22.2	18.24	20.0	14.70	20.4	13.63	21.6	13.19
1	0.08	12.73	24.9	17.67	22.5	14.23	22.9	13.17	24.3	12.73
1	0.09	12.29	27.5	17.12	24.9	13.80	25.3	12.74	26.8	12.29
1	0.10	11.86	30.0	16.59	27.3	13.37	27.6	12.31	29.3	11.86

Table 4.2: Quantile Hedging costs for a call option under borrowing constraint with parameters $S_0 = 100$, $K = 100$, $\mu = 0.08$, $\sigma = 0.3$.

T	Shortfall Prob	Initial Cost	% Gain	C_b 2	% Gain	C_b 5	% Gain	C_b 10	% Gain	C_b 1000
0.083	0	3.44		15.15		7.47		4.92		3.44
0.083	0.01	3.23	6.1	11.55	23.7	6.98	6.6	4.68	4.7	3.23
0.083	0.02	3.06	11.2	9.76	35.5	6.50	13.1	4.47	9.0	3.06
0.083	0.03	2.90	15.9	8.69	42.6	6.08	18.7	4.28	13.0	2.90
0.083	0.04	2.75	20.2	7.93	47.7	5.70	23.7	4.09	16.9	2.75
0.083	0.05	2.61	24.3	7.31	51.7	5.39	27.9	3.91	20.5	2.61
0.083	0.06	2.47	28.2	6.76	55.4	5.09	31.9	3.74	23.8	2.47
0.083	0.07	2.35	31.8	6.31	58.3	4.80	35.8	3.59	27.1	2.35
0.083	0.08	2.22	35.4	5.93	60.8	4.56	39.0	3.43	30.2	2.22
0.083	0.09	2.11	38.7	5.55	63.3	4.32	42.1	3.28	33.3	2.11
0.083	0.10	2.00	41.9	5.20	65.7	4.10	45.2	3.13	36.3	2.00
0.5	0	8.40		16.87		10.77		9.13		8.40
0.5	0.01	7.93	5.5	15.94	5.5	10.30	4.3	8.67	5.1	7.93
0.5	0.02	7.54	10.2	15.02	11.0	9.88	8.2	8.27	9.4	7.54
0.5	0.03	7.17	14.6	14.19	15.9	9.49	11.9	7.91	13.4	7.17
0.5	0.04	6.83	18.7	13.43	20.4	9.12	15.4	7.57	17.2	6.83
0.5	0.05	6.50	22.6	12.75	24.4	8.75	18.7	7.24	20.7	6.50
0.5	0.06	6.19	26.2	12.09	28.3	8.41	21.9	6.93	24.1	6.19
0.5	0.07	5.90	29.7	11.53	31.7	8.08	25.0	6.64	27.3	5.90
0.5	0.08	5.61	33.2	10.98	34.9	7.76	28.0	6.35	30.5	5.61
0.5	0.09	5.34	36.4	10.44	38.1	7.45	30.8	6.08	33.4	5.34
0.5	0.10	5.08	39.5	9.91	41.2	7.15	33.6	5.82	36.3	5.08
1	0	11.84		18.85		13.69		12.35		11.84
1	0.01	11.24	5.1	18.04	4.3	13.09	4.4	11.74	4.9	11.24
1	0.02	10.71	9.6	17.27	8.4	12.55	8.3	11.21	9.2	10.71
1	0.03	10.22	13.8	16.54	12.2	12.06	11.9	10.72	13.2	10.22
1	0.04	9.75	17.7	15.87	15.8	11.60	15.3	10.25	16.9	9.75
1	0.05	9.31	21.4	15.21	19.3	11.15	18.6	9.81	20.5	9.31
1	0.06	8.89	24.9	14.56	22.7	10.72	21.7	9.39	23.9	8.89
1	0.07	8.49	28.3	13.99	25.8	10.31	24.7	8.99	27.2	8.49
1	0.08	8.09	31.7	13.44	28.7	9.91	27.6	8.59	30.4	8.09
1	0.09	7.72	34.8	12.89	31.6	9.52	30.5	8.23	33.4	7.72
1	0.10	7.36	37.8	12.36	34.4	9.14	33.2	7.86	36.3	7.36

Table 4.3: Quantile Hedging costs for a call option under borrowing constraint with parameters $S_0 = 100$, $K = 110$, $\mu = 0.08$, $\sigma = 0.3$.

T	Shortfall Prob	Initial Cost	% Gain	C_b 2	% Gain	C_b 5	% Gain	C_b 10	% Gain	C_b 1000
0.083	0	0.61		12.52		4.64		2.01		0.61
0.083	0.01	0.48	21.0	6.73	46.2	3.55	23.6	1.72	14.4	0.48
0.083	0.02	0.39	36.1	4.58	63.4	2.73	41.2	1.44	28.4	0.39
0.083	0.03	0.32	48.3	3.48	72.2	2.16	53.4	1.22	39.4	0.32
0.083	0.04	0.25	58.6	2.68	78.6	1.77	61.9	1.00	50.1	0.25
0.083	0.05	0.20	67.4	2.17	82.7	1.43	69.3	0.83	58.8	0.20
0.083	0.06	0.16	74.6	1.66	86.7	1.12	75.8	0.69	65.6	0.16
0.083	0.07	0.12	80.7	1.33	89.4	0.93	80.1	0.56	71.9	0.12
0.083	0.08	0.08	86.4	1.08	91.4	0.74	84.1	0.44	78.3	0.08
0.083	0.09	0.06	90.0	0.81	93.5	0.55	88.1	0.33	83.6	0.06
0.083	0.10	0.04	93.3	0.58	95.3	0.39	91.5	0.22	88.8	0.04
0.5	0	4.72		13.97		7.30		5.47		4.72
0.5	0.01	4.32	8.5	12.74	8.8	6.84	6.3	5.07	7.3	4.32
0.5	0.02	3.98	15.6	11.51	17.6	6.43	11.9	4.73	13.4	3.98
0.5	0.03	3.69	21.9	10.50	24.8	6.05	17.1	4.44	18.9	3.69
0.5	0.04	3.41	27.7	9.60	31.3	5.68	22.2	4.16	23.9	3.41
0.5	0.05	3.16	33.1	8.88	36.5	5.32	27.0	3.91	28.5	3.16
0.5	0.06	2.92	38.0	8.16	41.6	5.01	31.4	3.66	33.0	2.92
0.5	0.07	2.70	42.7	7.51	46.3	4.71	35.5	3.44	37.1	2.70
0.5	0.08	2.49	47.3	7.00	49.9	4.42	39.4	3.22	41.2	2.49
0.5	0.09	2.30	51.3	6.49	53.5	4.15	43.2	3.00	45.1	2.30
0.5	0.10	2.11	55.2	5.97	57.3	3.87	46.9	2.80	48.7	2.11
1	0	8.11		15.76		10.03		8.64		8.11
1	0.01	7.55	6.9	14.86	5.7	9.47	5.6	8.09	6.5	7.55
1	0.02	7.07	12.8	14.01	11.1	8.99	10.4	7.61	12.0	7.07
1	0.03	6.64	18.1	13.18	16.4	8.54	14.9	7.18	17.0	6.64
1	0.04	6.23	23.1	12.45	21.0	8.11	19.1	6.77	21.7	6.23
1	0.05	5.85	27.8	11.77	25.3	7.71	23.1	6.39	26.1	5.85
1	0.06	5.50	32.2	11.08	29.7	7.33	26.9	6.03	30.2	5.50
1	0.07	5.16	36.4	10.44	33.8	6.96	30.6	5.70	34.1	5.16
1	0.08	4.83	40.5	9.88	37.3	6.62	34.0	5.37	37.9	4.83
1	0.09	4.53	44.1	9.35	40.7	6.28	37.4	5.07	41.4	4.53
1	0.10	4.23	47.8	8.82	44.0	5.94	40.7	4.77	44.8	4.23

Table 4.4: Quantile Hedging cost comparison to continuous-time minimal super-replication cost of call option under borrowing constraint with parameters $\mu = 0.08$, $\sigma = 0.3$.

S	K	T	min initial cost	C_b 2	5	10
100	90	0.083	$\hat{y}$	18.702	12.453	11.029
			$\tilde{y}$	18.699	12.449	11.024
100	90	0.5	$\hat{y}$	20.729	15.772	14.541
			$\tilde{y}$	20.721	15.759	14.522
100	90	1	$\hat{y}$	22.815	18.489	17.437
			$\tilde{y}$	22.805	18.462	17.397
100	100	0.083	$\hat{y}$	15.148	7.477	4.923
			$\tilde{y}$	15.146	7.473	4.916
100	100	0.5	$\hat{y}$	16.886	10.805	9.172
			$\tilde{y}$	16.869	10.770	9.134
100	100	1	$\hat{y}$	18.872	13.717	12.430
			$\tilde{y}$	18.846	13.693	12.345
100	110	0.083	$\hat{y}$	12.519	4.649	2.016
			$\tilde{y}$	12.518	4.642	2.006
100	110	0.5	$\hat{y}$	13.981	7.329	5.517
			$\tilde{y}$	13.974	7.297	5.469
100	110	1	$\hat{y}$	15.778	10.094	8.679
			$\tilde{y}$	15.755	10.028	8.644

Table 4.5: Quantile Hedging costs for an up-and-out barrier call option with parameters $S_0 = 100$, $B = 130$, $\mu = 0.08$, $\sigma = 0.3$.

T	Shortfall Prob	initial cost for $K = 100$	% Gain	initial cost for $K = 90$	% Gain
0.0833	0	3.39		10.37	
0.0833	0.01	3.19	5.8	10.09	2.7
0.0833	0.02	3.02	10.9	9.83	5.2
0.0833	0.03	2.86	15.6	9.59	7.5
0.0833	0.04	2.72	19.9	9.35	9.8
0.0833	0.05	2.58	24.0	9.12	12.0
0.0833	0.06	2.45	27.9	8.91	14.1
0.0833	0.07	2.32	31.6	8.69	16.2
0.0833	0.08	2.20	35.1	8.48	18.2
0.0833	0.09	2.09	38.4	8.28	20.1
0.0833	0.1	1.98	41.7	8.08	22.1
0.5	0	2.95		6.80	
0.5	0.01	2.73	7.5	6.50	4.4
0.5	0.02	2.53	14.1	6.22	8.5
0.5	0.03	2.34	20.6	5.95	12.5
0.5	0.04	2.16	26.7	5.68	16.4
0.5	0.05	2.00	32.1	5.44	20.0
0.5	0.06	1.85	37.3	5.20	23.6
0.5	0.07	1.69	42.5	4.96	27.1
0.5	0.08	1.54	47.7	4.72	30.6
0.5	0.09	1.41	52.1	4.50	33.8
0.5	0.1	1.30	56.1	4.29	36.9
1	0	1.67		3.93	
1	0.01	1.46	12.8	3.64	7.5
1	0.02	1.28	23.5	3.37	14.2
1	0.03	1.11	33.7	3.12	20.6
1	0.04	0.95	43.5	2.87	27.0
1	0.05	0.82	51.0	2.66	32.4
1	0.06	0.70	58.1	2.45	37.7
1	0.07	0.58	65.1	2.24	43.0
1	0.08	0.47	72.0	2.04	48.1
1	0.09	0.37	77.8	1.85	53.0
1	0.1	0.31	81.7	1.69	57.0

Bibliography

[1] Bensaid, B. and Lesne, J. P., Pagés, H., and Scheinkman, J. (1992). Derivative asset pricing with transaction costs. *Mathematical Finance* , 2, 63–86.

[2] DeGennaro, R.P. and Robotti, C. (2007). Financial Market Frictions. *Economic Review, Federal Reserve Bank of Atlanta*, 92, 1–16.

[3] Hodges, R. and Neuberger, K. (1989). Optimal replication of contingent claims under transaction costs. *Review Futures Markets* , 222–239.

[4] Melnikov, A. and Skornyakova, V. (2005). Quantile hedging and its application to life insurance. *Statistics and Decisions*, 4(23), 301-316.

[5] Cochrane, J. and SaáRequejo, J. (2000). Beyond arbitrage: good-deal asset price bounds in incomplete markets. *Journal of Political Economy*, 108, 79–119.

[6] Musiela, M. and Zariphopoulou, T. (2004). An example of indifference prices under exponential preferences. *Finance and Stochastics* , 8, 229–239.

[7] Black, F. and Scholes, M. (1973). The valuation of options and corporate liability. *Journal of Political Economy*, 81, 637–654.

[8] Carassus, L. and Rásonyi, M. (2007). Convergence of Utility Indifference Prices to the Superreplication Price: the Whole Real Line Case . *Acta Applicandae Mathematicae*, 96, 119–135.

[9] Rubinstein, M. (1992). "Guiding Force" in From Black-Scholes to Black Holes: New Frontiers in Options. *Risk Magazine.*

[10] Föllmer, H. and Leukert, P. (1999). Quantile hedging. *Finance and Stochastics*, 3, 251–273.

[11] Bouchard, B., Elie, R. and Touzi, N. (2009). Stochastic Target Problems with Controlled Loss. *SIAM Journal on Control and Optimization*, 48(5), 3123–3150.

[12] El Karoui, N. and Quenez, M. C. (1995). Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control and Optimization*, 33(1), 29-66.

[13] Naik, V. and Uppal, R. (1994). Leverage Constraints and the Optimal Hedging of Stock and Bond Options. *Journal of Financial and Quantitative Analysis*, 29(2), 199–222.

[14] Broadie, M., Cvitanic, J, and Soner, H. M. (1998). Optimal replication of contingent claims under portfolio constraints. *Review of Financial Studies*, 11, 59–79.

[15] Soner, H. M. and Touzi, N. (2002). Stochastic Target Problems, dynamic programming and viscosity solutions. *SIAM Journal on Control and Optimization*, 41, 404-424.

[16] Spivak, G.and Cvitanic, J. (1999). Maximizing The Probability of a Perfect Hedge. *Annals of Applied Probability*, 9(4), 1303.

[17] Lindberg, P. (2010). Optimal partial hedging in a discrete-time market as a knapsack problem. *Mathematical Methods of Operations Research*, 72(3), 433 - 451.

[18] Merton, R. C. (1973). Theory of Rational Option Pricing. *Bell Journal of Ecnonomics*, 4(1), 141-183.

[19] Krutchenko, R. N. and Melnikov, A. (2001). Quantile hedging for a jump-diffusion financial market. *Trends in Mathematics*

[20] Fleming, W. and Soner, H. M. (1993). Controlled Markov Processes and Viscosity Solutions. *Springer Verlag.*

[21] Perez-Hernandez, L. (2007). On the existence of an efficient hedge for an American contingent claim within a discrete time market. *Quantitative Finance*, 7(5), 547-551.

[22] Pınar, M. Ç. (2011). Buyer's quantile hedge portfolios in discrete-time trading. *Quantitative Finance*, forthcoming, 1-10.

[23] Bernardo, A.E. and Ledoit, O. (2000). Gain, loss, and asset-pricing. *Journal of Political Economy* , 108, 144–172.

[24] Pınar, M. Ç., Salih, A. and Camci, A. (2007). Expected gain-loss pricing and hedging of contingent claims in incomplete markets by linear programming. *European Journal of Operational Research*, 201, 770–785.

[25] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent Measures of Risk. *European Journal of Operational Research*, 9, 203–228.

Chapter 5

Optimal Dividend Policy with Random Interest Rates

Erdinç Akyıldırım, İ. Ethem Güney, Jean-Charles Rochet, and H. Mete Soner

5.1 Introduction

Since Jeanblanc-Picqué & Shiryaev [8] and Radner & Shepp [10], a sizable literature has investigated the optimal dividend policy problem for a company that is not allowed to issue new securities or obtain a new loan from a bank. The default time is then defined as the first time when the cash reserves of the company fall below zero. In that case, the optimal dividend policy is simple and natural: distribute dividends whenever the level of cash reserves exceeds a certain threshold that depends on the characteristics (drift, volatility) of the cash flow process and the interest rate demanded by shareholders.

An interesting extension of this problem is to investigate how the optimal dividend policy is modified when the profitability of the firm changes over time, due in particular to business cycle fluctuations. As clearly shown by Gertler & Hubbard [5] and more recently by Hackbart, Miao and Morellec [6], macroeconomic conditions have indeed a strong impact on dividend policies through the changes in the profitability of individual firms that they induce. For example, Cadenillas & Sotomayor [2] solve for the optimal dividend policy when the drift and the volatility of the cash flow process are governed by a Markov chain representing macroeconomic fluctuations. Bolton, Chen & Wang [1] study more generally the impact of changing macroeconomic conditions on both the financial and investment policies of the firms. However, Gertler & Hubbard [5] also show that macroeconomic conditions directly influence payments to shareholders, even independently of each firm's specific earnings performance. Two natural channels for this influence are the fluctuations in interest rates demanded by investors, and the conditions of the credit market.

The purpose of this paper is to examine how these macroeconomic fluctuations influence the dividend policies of firms, even in the absence of fluctuations in their earning processes. In other words, we study the polar case to the one considered in the literature: the drift and volatility of the cash flow process are constant, but the interest rate demanded by investors follows a Markov chain.[1]

Section §5.2 presents the model and the mathematical characterization of the optimal dividend policy (Theorem 5.2.1). In the next section, we establish several important properties of the value function. In subsection §5.3.1, we show that the value function remains concave in the level of cash holdings, even when interest rates are stochastic (Theorem 5.3.3). The concavity of the value function allows us to prove that it is a smooth solution of the corresponding dynamic programming equation (Proposition 5.3.4). In particular, it satisfies the *smooth fit* condition which is crucial in the numerical resolution of these types of problems. These mathematical results are necessary to establish the fundamental economic result of the paper in subsection §5.3.3: the firm will distribute dividends more often when interest rates are high than when they are low (Proposition 5.3.5). This result comes from the fact that the opportunity cost of cash reserves is higher when the interest rates demanded by investors are high. However, it does not fit well with the empirical evidence, given that firms actually tend to distribute less dividends during recessions (when interest rates are high) than during booms (when interest rates are low) even when the changes in firms' individual profitability are corrected for (Gertler

[1] In a recent paper, Jiang and Pistorius [9] consider the case where both the profitability of the firm and the discount factor follow a Markov chain. However they do not allow the firm to issue new equity.

& Hubbard [5]). This suggests that other macroeconomic factors, such as the size of frictions on financial markets, must play a role. This is why section §5.4 introduces the possibility for the firm to make new equity issuances. When the cost of these new issues (a proxy for the size of financial frictions) is substantially higher during recessions than during booms, the ranking of dividend thresholds is reversed, and firms now distribute more dividends during booms than during recessions.

We also provide numerical evidence for the above conclusions. In particular, in subsection §5.3.4, the sensitivity analysis with respect to mean and volatility of the cash flow rate and jump rates between two different interest rate regimes are presented. The mathematical results proved in Section §5.3 are also essential in constructing and verifying the numerical algorithm. Section §5.4 gives several numerical illustrations of the case where new equity issuance is possible.

5.2 Model and Characterization of the Solution

Uncertainty is described by $(\Omega, \mathbb{F}, \mathbb{P})$, a filtered probability space satisfying the usual assumptions[2]. Let B_t be a one-dimensional standard Brownian motion and $\{i_t\}_{t\geq 0}$ be a simple stationary Markov process taking values in $\{0, 1\}$ with jump rates $\lambda(0), \lambda(1) > 0$. The process $\{i_t\}_{t\geq 0}$ is assumed to be independent from the Brownian motion. The state $i = 0$ is the "good" economic state with a lower interest rate $r_\ell > 0$ and $i = 1$ corresponds to the "bad" state with interest rate $r_h > r_\ell > 0$. We also set $\lambda_\ell := \lambda(0)$ and $\lambda_h := \lambda(1)$.

The cash holdings $\{X_t\}_{t\geq 0}$ of the company follow a diffusion process. Positive dividend payments of any size can be made at any time. However, the cash level is supposed to remain nonnegative at all times. This constraint clearly places a restriction of the possible dividend size. Mathematically,

$$dX_t = \mu dt + \sigma dB_t - dL_t, \tag{5.2.1}$$

where $\mu, \sigma > 0$ are given constants and the *cumulative dividend payments* L_t is an adaptive, nondecreasing, càdlàg process with $L_{0^-} = 0$. Given a dividend process L and an initial condition $x \in \mathbb{R}$, let $X^{x,L}$ be the unique solution of (5.2.1), i.e.,

[2]See [7] for details.

$$X_t^{x,L} = x + \mu t + \sigma B_t - L_t, \quad t \geq 0.$$

Let $\theta = \theta^{x,L}$ be the first exit time of $X^{x,L}$ from the positive real line. This variable θ defines the time of bankruptcy. In what follows we will suppress the dependence on x, L unless this dependence is important. We say that L is *admissible* at the initial level x, if $X_t^{x,L} \geq 0$, for all time $t \in [0, \theta^{x,L}]$ with probability one. We denote the set of all admissible strategies by $\mathcal{A}(x)$. We note that the admissibility condition is relevant only at the exit time. Indeed, we only require that the cash level process does not jump into negative real line. In economic terms, this means that shareholders can never distribute themselves a dividend that exceeds the cash holdings of the firm. Hence, $X_\theta^{x,L} = 0$. Since the dividend policy beyond the exit time is irrelevant, we simply set $L_t = L_\theta$ for all $t \geq \theta$. In particular, $L_\theta - L_{\theta^-} = X_{\theta^-}$.

The optimal dividend problem is to maximize

$$J(x,i,L) := \mathbb{E}\left[\int_0^\theta \Lambda_t \, dL_t \;\middle|\; i_0 = i, X_{0^-} = x\right], \quad \Lambda_t := \exp\left(-\int_0^t r(i_u)du\right).$$

The *value function* is then defined by

$$v(x,i) := \sup_{\mathcal{A}(x)} J(x,i,L), \quad v_\ell(x) := v(x,0), \quad v_h(x) := v(x,1). \tag{5.2.2}$$

The case of a deterministic (and constant) interest rate (i.e., $r_\ell = r_h$) is exactly the problem studied by Picqué-Jeanblanc & Shiryayev [8] and Radner & Shepp [10]. For future reference, we record that the value function with constant interest rate r is given by

$$V(x,r) := \sup_{L \in \mathcal{A}(x)} \mathbb{E}\left[\int_0^\theta e^{-rt} dL_t | X_{0^-} = x\right]. \tag{5.2.3}$$

Then, it is clear that

$$0 \leq V(x,r_h) \leq v_h(x) \leq v_\ell(x) \leq V(x,r_l), \qquad \forall\, x \in \mathbb{R}^+. \tag{5.2.4}$$

5.2.1 Characterization of the Solution

Our main mathematical result is the following characterization of the value function. The existence part of this theorem will be proved in several steps in the subsequent sections. The uniqueness follows from the classical verification argument (see for instance [4]). This characterization of the value function and the properties of the thresholds are essential in our numerical experiments. Indeed, the numerical algorithm is based on these properties. Moreover, the uniqueness ensures that the computed functions are in fact equal to the value function.

Theorem 5.2.1. The value function $v = (v(\cdot, 0), v(\cdot, 1)) = (v_h, v_\ell)$ is the unique concave function satisfying the following conditions:

- $v_\ell, v_h \in C^2([0, \infty)$ and $v_l(0) = v_h(0) = 0$;
- $v'(x, i) \geq 1$ for all x;
- For every $x > 0$ and $i \in \{0, 1\}$, $r(i)v(x, i) - \mathcal{L}v(x, i) \geq 0$, where

$$\mathcal{L}v(x, i) := \mu v'(x, i) + \frac{\sigma^2}{2} v''(x, i) + \lambda(i)[v(x, i + 1) - v(x, i)]; \tag{5.2.5}$$

  with the convention that $i + 1$ denotes the other state than i.

- There are two positive thresholds $0 < x_h := x(1)$ *and* $x_\ell := x(0) < \infty$ such that

$$v'(x, i) = 1, \quad \text{for } x \geq x(i), \quad \text{and} \quad r(i)v(x, i) - \mathcal{L}v(x, i) = 0, \quad \text{for } x \leq x(i).$$

The above characterization of the value function also provides the structure of the optimal dividend policy. The optimal dividend policy is simple: only distribute dividends when cash holdings exceed threshold $x(i)$, which depends on the state i of the economy. This is done exactly as in the deterministic interest rate case. Namely, if the initial cash holdings x exceed $x(i)$, then an initial dividend of $x - x(i)$ is distributed. In later times, dividends are paid only when the cash holdings reach $x(i)$ again. When the state of the economy changes from good to bad (equivalently when i jumps from zero to one), then cash holdings may be larger than $x(1)$ and a dividend payment of the difference is optimal. Then, one proceeds as before.

The above theorem also proves that the value function is a classical solution of the dynamic programming equation,

$$\min\left\{\ r(i)v(x,i)-\mathcal{L}v(x,i)\ ,\ v'(x,i)-1\ \right\}=0, \qquad x>0,\ i=1,2, \tag{5.2.6}$$

together with boundary condition $v(0,i)=0$.

5.2.2 Elementary Properties

In this subsection, we prove several simple properties.

Lemma 5.2.2. The value function v is Lipschitz continuous at the origin and

$$v(0,i)=0, \quad v(x+y,i)\geq v(x,i)+y, \quad \forall\ x,y\geq 0,\ i=0,1.$$

Proof. Since σ is not null, the only admissible process at $x=0$ is $L=0$. This proves that $v(0,i)=0$. We also emphasize that at time zero, L^y has a jump of size at least y. Also, for any given (x,y) and $L\in\mathcal{A}(x)$, we set $L^y_t:=L_t+y$ for $t\geq 0$ (with, as it is required $L^y_{0^-}=0$).

Then, if one starts with cash holdings $x+y$ at $t=0$ and uses the dividend policy L^y, cash holdings are characterized by $\{\hat{X}_t\}_{t\geq 0}$ defined by

$$\begin{aligned}\hat{X}_t \ &:= \ X_t^{x+y,L^y}=x+y+\mu t+\sigma W_t-L^y_t\\ &= \ x+\mu t+\sigma W_t-L_t=X_t^{x,L}=:X_t,\end{aligned}$$

for all $t\geq 0$. In particular, the exit time $\hat{\theta}$ of $\hat{X}$ from $(0,\infty)$ is the same as that of X. Hence,

$$v(x+y,i)\geq J(x+y,i,L^y)=\mathbb{E}\left[\int_0^\theta \Lambda_t dL^y_t\right]=y+\mathbb{E}\left[\int_0^\theta \Lambda_t dL_t\right].$$

Since $L\in\mathcal{A}(x)$ is arbitrary,

$$v(x+y,i)\geq y+v(x,i), \quad \forall\ (x,y)\in\mathbb{R}^+,\ i=0,1.$$

Recall the deterministic value function defined in (5.2.3) and the inequality (5.2.4). Hence for any $x \geq 0$ and i,

$$V(0, r_\ell) = v(0, i) = 0 \leq v(x, i) \leq V(x, r_\ell).$$

The function V is known explicitly (see [8]) and it is Lipschitz continuous. Hence, v is Lipschitz continuous at the origin, i.e., there is a constant K such that

$$0 = v(0, i) \leq v(x, i) \leq Kx$$

for all $x \geq 0$. □

In this context, the standard dynamic programming principle states that for any initial point (x, i) and any stopping time $\tau \leq \theta$,

$$v(x, i) = \sup_{L \in \mathcal{A}(x)} \mathbb{E}\left[\int_0^\tau \Lambda_t dL_t + \Lambda_\tau \; v\left(X_\tau^{x,L}, i_\tau\right)\right]. \tag{5.2.7}$$

Our next result, is a step towards proving the concavity of the value function. Indeed, the concavity is equivalent to the condition (5.2.8) below with $c_0 = 0$.

Lemma 5.2.3. There exists a constant $c_0 > 0$ such that for all $0 \leq x < y$ and $i \in \{0, 1\}$,

$$v(x, i) + v(y, i) - 2v((x + y)/2, i) \leq c_0. \tag{5.2.8}$$

Proof. Recall the value function defined in (5.2.3) and the inequality (5.2.4). Then,

$$v(x, i) + v(y, i) - 2v((x + y)/2, i) \leq V(y, r_\ell) + V(x, r_\ell) - 2V((x + y)/2, r_h).$$

The function V is known explicitly and such that there exists a constant $c(r) > 0$ so that

$$x \leq V(x, r) \leq c(r) + x, \quad \forall \; x, r > 0.$$

We now combine the two inequalities to obtain,

$$v(x,i)+v(y,i)-2v((x+y)/2,i)\leq[c(r_\ell)+x]+[c(r_\ell)+y]-2((x+y)/2)\leq 2c(r_\ell).$$

□

Indeed, the viscosity property is proved exactly as in Theorem 5.1, page 311 in [4]. Moreover, the uniqueness of this solution can be proved by the techniques developed in [4]. But this result is not needed in this paper.

Lemma 5.2.4. The value function is a continuous viscosity solution of the dynamic programming equation (5.2.6).

5.3 Value Function

In this section, we establish several important properties of the value function.

5.3.1 Concavity

In this section, we prove that the value function is concave. We start by showing this in an interval near the origin.

Lemma 5.3.1. There exists $x_0>0$ such that for both $i=0,1$,

$$-v''(\cdot,i)\geq 0, \quad \text{on } (0,x_0),$$

in the viscosity sense.

Proof. We first choose $x_0>0$ so that

$$|r(i)v(x,i)-\lambda(i)[v(x,i+1)-v(x,i)]|\leq\mu, \quad \forall\ x\in[0,x_0], i\in\{0,1\}.$$

This is possible as v is continuous at the origin with value zero.

We need to show that for $\varphi(.,i)\in C^2(\mathbb{R})$ for each i, which depends on the state of the economy i, if

$$(v-\varphi)(x^*,i)=localmin(v-\varphi)(\cdot,i)$$

at some $x^* \in (0, x_0)$, then $\varphi''(x^*) \leq 0$.

Indeed, let φ be as above. Then, by the viscosity supersolution property of v we have

$$r(i)v(x^*, i) - \mu\varphi'(x^*) - \frac{\sigma^2}{2}\varphi''(x^*) - \lambda(i)[v(x^*, i+1) - v(x^*, i)] \geq 0,$$

and $\varphi'(x^*) \geq 1$. Hence,

$$-\varphi''(x^*) \geq \frac{1}{\sigma^2}\left(-r(i)v(x^*, i) + \mu + \lambda(i)[v(x^*, i+1) - v(x^*, i)]\right).$$

By the choice of x_0, the right hand side of the above inequality is non-negative. Therefore, $-\varphi'' \geq 0$. □

The following is an immediate corollary of the above Lemma.

Corollary 5.3.2. There exists $x^* > 0$ such that $v(\cdot, i)$ is concave on $[0, x^*]$ and

$$v'(x, i) \geq v'(x^*, i) > 1, \quad \forall\, i \in \{0, 1\},\ x \in [0, x^*].$$

Proof: The concavity of v near the origin follows from the previous results and the theory of viscosity solutions. Also

$$v(h, i) = v(h, i) - v(0, i) \geq V(h, r_h) > (1 + \delta)h,$$

for some $\delta > 0$. Hence, $v'(0, i) \geq 1 + \delta$. Set

$$x^* = \sup\{x :\ v(\cdot, i) \textit{ is concave on } [0, x] \textit{ and } v'(x, i) \geq 1 + \delta/2\}.$$

Then, it is clear that $x^* > 0$. □

The following is proved in the Appendix A.

Theorem 5.3.3. $v(\cdot, i)$ is concave for $i \in \{0, 1\}$.

5.3.2 Smooth Fit

In this section, we use the concavity of the value function to show that it is twice continuously differentiable. This statement is equivalent to the smooth fit property at the thresholds. The smoothness of the value function immediately implies that it is a classical solution of the dynamic programming equation (5.2.6).

Proposition 5.3.4 (Smooth Fit)**.** The value function is twice continuously differentiable in the x variable.

Proof. Set

$$x(i) = inf\{x : \ 1 \in \partial v(x,i)\}, \quad i = 0, 1 \tag{5.3.1}$$

where $\partial v(x,i)$ denotes the subdifferential of $v(\cdot, i)$ at x (we refer reader to [11] for the definition and the properties of subdifferentials of convex functions). By Lemma 5.2.2 $x(i) > 0$. Also, since $v' \geq 1$ in the viscosity sense, concavity of v implies,

$$v'(x,i) = 1, \quad \forall\ x \geq x(i), \qquad \text{and} \qquad v'(x,i) > 1, \quad \forall\ x \in [0, x(i)).$$

Then, since v satisfies the dynamic programming equation (5.2.6),

$$r(i)v(x,i) - \mathcal{L}v(x,i) = 0 \quad \forall x \in (0, x(i)),$$

the elliptic regularity implies that

$$v(\cdot, i) \in C^{\infty}((0, x(i))).$$

Step 1. First, we show that $\partial v(x(i), i) = \{1\}$.

Suppose to the contrary that

$$\partial v(x(i), i) = [1, p]$$

for some $p > 1$. Then, for any $\varepsilon > 0$, it is straightforward to construct a smooth test function φ_ε so that

$$\sup(v(\cdot,i)-\varphi_\varepsilon(\cdot)) = v(x(i),i)-\varphi_\varepsilon(x(i)) = 0,$$

$\varphi_\varepsilon''(x(i)) = -1/\varepsilon$ and $\varphi_\varepsilon'(x(i)) \in (1,p)$. The viscosity property of $v(\cdot,i)$ implies that

$$r(i)v(x(i),i) - \mu\varphi_\varepsilon'(x(i)) - \frac{\sigma^2}{2}\varphi_\varepsilon''(x(i)) - \lambda(i)[v(x(i),i+1) - v(x(i),i)] \le 0.$$

For $\varepsilon > 0$ sufficiently small, this is a contradiction. Hence, $\partial v(x(i),i)$ is a singleton $\{1\}$ and $v \in C^1([0,\infty))$.

Step 2. We now show that $v \in C^2$.

The only point at which v may not be twice differentiable is $x(i)$ and

$$v''(x,i) = 0, \;\; \forall x > x(i).$$

Set

$$\gamma = \liminf_{x\uparrow x(i)} v''(x,i).$$

Then there exists $x_n < x(i)$ converging to $x(i)$, so that $v''(x_n,i) \to \gamma$. By the first step, $v'(x_n,i) \to 1$. Moreover, the elliptic equation holds at all x_n's. Hence,

$$\begin{aligned} r(i)v(x(i),i) - \mu - \frac{\sigma^2}{2}\gamma - \lambda(i)[v(x(i),i+1) - v(x(i),i)] \\ = \lim_{n\to\infty} r(i)v(x_n,i) - \mathcal{L}v(x_n,i) = 0. \end{aligned} \tag{5.3.2}$$

The dynamic programming equation (5.2.6) implies that at any $x > x(i)$,

$$0 \le r(i)v(x,i) - \mathcal{L}v(x,i) = r(i)v(x,i) - \mu - \lambda(i)[v(x,i+1) - v(x,i)].$$

Hence as $x \downarrow x(i)$

$$r(i)v(x(i),i) - \mu - \lambda(i)[v(x(i),i+1) - v(x(i),i)] \ge 0.$$

The above inequality, together with (5.3.2) imply that $\gamma \ge 0$. However, by concavity,

$v'' \leq 0$. Hence, $\gamma = 0$ and

$$0 \leq \liminf_{x \uparrow x(i)} v''(x,i) \leq \limsup_{x \uparrow x(i)} v''(x,i) \leq 0.$$

Therefore, v is twice differentiable at $x(i)$. $\square$

5.3.3 Dividend Thresholds

In the previous sections, we have shown that v is a concave, twice continuously differentiable, classical solution of (5.2.6). By concavity and Lemma 5.2.2, there are $x(i) > 0$, $i = 0, 1$ such that

$$v'(x,i) = 1 \quad \text{for} \quad x \geq x(i), \quad \text{and} \quad v'(x,i) > 1, \; r(i)v(x,i) - \mathcal{L}v(x,i) = 0, \quad \text{on} \quad [0, x(i)).$$

Indeed,

$$x(i) := inf\{x : \; v'(x,i) = 1 \;\}, \quad \text{and} \quad x_\ell := x(0), \;\; x_h := x(1).$$

The following is proved in Appendix A.

Proposition 5.3.5. Let $x_\ell, x_h > 0$ be as above. Then, $x_\ell \geq x_h$.

5.3.4 Sensitivity Analysis

In this section we give numerical illustrations of the value function and the sensitivities of the dividend thresholds with respect to mean and volatility of the cash flow process and the jump rate between low and high interest rate regimes.

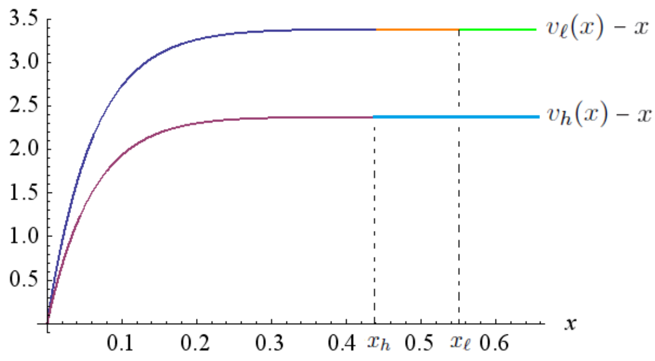

Figure 5.1: Value function with parameters $\mu = 0.18,\ \sigma = 0.15,\ \lambda = 0.1,\ r_l = 0.02,\ r_h = 0.1,\ x_h = 0.4386,\ x_l = 0.5528.$

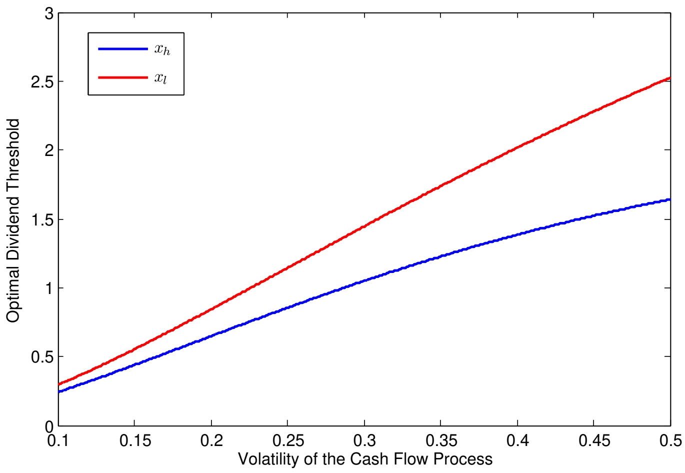

Figure 5.2: Sensitivities of x_h and x_l wrt σ with parameters $\mu = 0.18,\ \lambda = 0.1,\ r_l = 0.02,\ r_h = 0.1.$

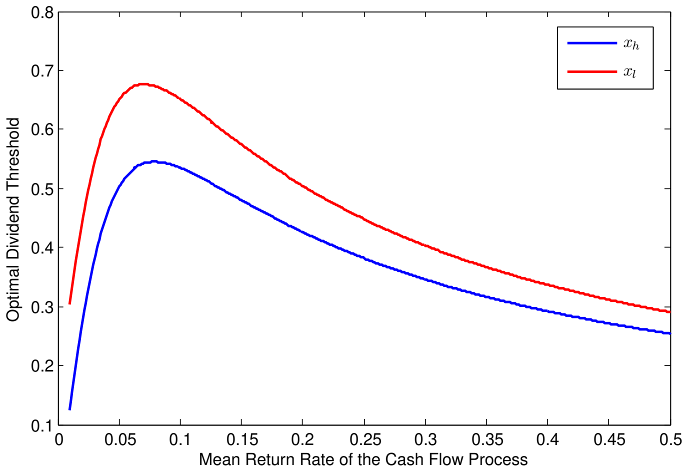

Figure 5.3: Sensitivities of x_h and x_l wrt μ with parameters $\sigma = 0.15$, $\lambda = 0.1$, $r_l = 0.02$, $r_h = 0.1$.

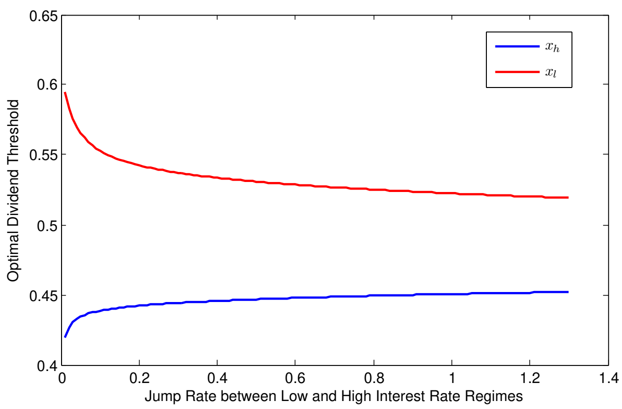

Figure 5.4: Sensitivities of x_h and x_l wrt μ with parameters $\mu = 0.18$, $\sigma = 0.15$, $r_l = 0.02$, $r_h = 0.1$.

5.4 Issuance

In this section, we enlarge the set of financial policies available to the firm, by allowing it to issue new shares, in addition to distribute dividends. Using the previous notation, the cash level process is now given by

$$X_t = x + \mu t + \sigma B_t - L_t + I_t, \tag{5.4.1}$$

where I_t is the total amount of cash raised up to time t (cumulated issuance process, net of issuance costs). We assume[3] that I is piece-wise constant and has then form

$$I_t = \sum_{k=1}^{\infty} \xi_k \chi_{\{t \geq \tau_k\}}, \tag{5.4.2}$$

where $0 \leq \tau_1 < \ldots < \tau_k < \tau_{k+1}$ are stopping times at which equity issues are made and $\xi_k \geq 0$ are the issuance sizes. Then, the optimization problem that the firm faces is to maximize[4]

$$J(x, i, L, I) := \mathbb{E}\left[\int_0^\theta \Lambda_t dL_t - \sum_{k=1}^{\infty} \Lambda_{\tau_k}\left(\xi_k + \gamma(i_{\tau_k})\right) \;\Big|\; i_0 = i, X_{0^-} = x\right], \tag{5.4.3}$$

where $\gamma(i) > 0$ is the fixed cost of issuance when the economy is in state i. The interpretation of functional J is straightforward. Since there is a fixed cost $\gamma(i)$ of issuance (which depends on the state i of the economy), new issues will be lumpy and occur at discrete times τ_1, τ_2,... . Since there is no marginal cost of issuance, the total amount of cash raised at date τ_k is just $\xi_k + \gamma(i_{\tau_k})$. Functional J represents expected present value of future dividend payments, net of equity issuances, as in [3].

The value function

$$v(x, i) := \sup_{L, I \in \mathcal{A}(x)} J(x, i, L, I)$$

[3]Given the presence of a fixed issuance cost, such a policy is indeed optimal without loss of generality.
[4]See [3] for a discussion of the objective function.

is the unique viscosity solution of

$$\begin{aligned}\min\Big\{ r(i)v(x,i) - \mathcal{L}v(x,i)\ ;\ v'(x,i) - 1\ ; \\ v(x,i) - \sup_{\xi\geq 0}\left(v(x+\xi,i) - \xi - \gamma(i)\right)\Big\} = 0.\end{aligned} \tag{5.4.4}$$

We distinguish the cases when the cost structure depends on the point process and when not.

5.4.1 Constant Issuance Cost

The following lemma shows that when $\gamma(i) \equiv \gamma$, it is never optimal to issue new equity before the cash reserves are zero. This is consistent with the results of [3] in the case where interest rates are constant.

Lemma 5.4.1. Suppose γ is independent of i. Then, it is never optimal to issue new equity when the cash level is non zero. Hence, v is the unique solution of

$$\min\left\{ r(i)v(x,i) - \mathcal{L}v(x,i)\ ;\ v'(x,i) - 1\ \right\} = 0,$$

with boundary condition

$$v(0,i) = \max\{0\ ;\ \sup_{\xi\geq 0}\left(v(\xi,i) - \xi - \gamma\right)\}.$$

Moreover for any $x > 0$,

$$v(x,i) > \sup_{\xi\geq 0}\left(v(x+\xi,i) - \xi - \gamma\right).$$

Proof.

Fix $x \geq 0$ and let $(L, I) \in \mathcal{A}(x)$ be any admissible dividend-issuance policy. Then, I is as in (5.4.2). Suppose that $X_{\tau_1} > 0$. Define $\tilde{I}$ simply by removing the first issuance, i.e.,

$$\tilde{I}_t = \sum_{k=2}^{\infty} \xi_k \chi_{\{t \geq \tau_k\}} = I_t - \xi_1 \chi_{\{t \geq \tau_1\}}.$$

The new strategy $(L, \tilde{I})$ may not be admissible, but the corresponding cash flow process $\tilde{X}$ exists and is given by

$$\tilde{X}_t = x + \mu t + \sigma B_t - L_t + \tilde{I}_t.$$

Set

$$\tau := \inf\{t \geq \tau_1 \ : \ \tilde{X}_t \leq 0\},$$

or infinity, if the above set is empty. Since we have assumed that $X_{\tau_1} > 0$, $\tau > \tau_1$.

We now define another issuance strategy $\hat{I}$ by

$$\hat{I}_t = \tilde{I}_t + \xi_1 \chi_{\{t \geq \tau\}}.$$

Then, it is clear that $\hat{I}_t = I_t$ for all $t \geq \tau$. Let $\hat{X}$ be the corresponding cash level process, i.e.,

$$\tilde{X}_t = x + \mu t + \sigma B_t - L_t + \hat{I}_t.$$

Then,

$$\hat{X}_t = \begin{cases} \tilde{X}_t, & \text{for } t \in [0, \tau), \\ X_t, & \text{for } t \geq \tau. \end{cases}$$

The above characterization of $\hat{X}$ shows that $\hat{X}_t \geq 0$ for all $t \geq 0$. Hence, $(L, \hat{I})$ is indeed admissible. Moreover,

$$J(x, i, L, \hat{I}) = J(x, i, L, I) + \mathbb{E}\left[(\Lambda_{\tau_1} - \Lambda_\tau)\, \xi_1\right] > J(x, i, L, I),$$

where the final inequality follows from the fact that $\tau > \tau_1$.

The above argument shows that it is enough to consider issuance strategies for which

$X_{\tau_1} = 0$. By induction we can show that this result extends to all issuance times and we need only to consider strategies with $X_{\tau_k} = 0$ for every k. This is exactly the statement of the Lemma. $\square$

5.4.2 Issuance with Random Costs

If the cost structure γ depends on i, then the above result no longer holds. This is illustrated in the following numerical example where $\gamma(1)$ is much larger than $\gamma(0)$. We use the following parameter values:

$$\mu = 0.18,\ \sigma = 0.5,\ \lambda = 0.1,\ r(0) = 0.02,\ r(1) = 0.1.$$

For this set of parameter values the value function is twice continuously differentiable except one point, x_I, and has the following form. There are thresholds $0 < x_I < x_\ell < x_h$. Set

$$\text{Region 1} := (0, x_I), \quad \text{Region 2} := (x_I, x_\ell), \quad \text{Region 3} := (x_\ell, x_h).$$

In region 1, the firm issues new equity when the interest rate is low (but not when it is high). The two other regions are associated with dividend thresholds x_l and x_h like before. Thus, the value function satisfies

$$\begin{aligned}
&v(x,0) = v(x_\ell,0) - (x_\ell - x) - \gamma(0), \quad x \in \text{Region 1},\\
&r(0)v(x,0) = \mathcal{L}v(x,0), \quad x \in \text{Region 2},\\
&v'(x,0) = 1, \quad x \geq x_\ell,\\
&r(1)v(x,1) = \mathcal{L}v(x,1), \quad x \leq x_h,\\
&v'(x,1) = 1, \quad x \geq x_h.
\end{aligned}$$

Therefore the optimal strategy is given as follows. The fixed cost $\gamma(1)$ is so high that it is never optimal to issue new equity if the state i is equal to one (equivalently, if the interest rate is high). The dividend threshold for $r = r_h$ is x_h and when $r = r_l$ it is x_ℓ. Interestingly, $x_\ell < x_h$ while without issuance the opposite inequality always holds, c.f, Proposition (5.3.5). For $i = 0$, if the cash level is sufficiently small, i.e., if in Region 1, then the firm issues new equity. In Region 2, the firm does not take any action and pays dividends when $x > x_\ell$. The value function is shown in the figure below, for the parameter values

$$\gamma(0) = 0.48, \quad r(0) = 0.02, \quad r(1) = 0.1, \quad \lambda = 0.1, \quad \sigma = 0.5, \quad \mu = 0.18. \tag{5.4.5}$$

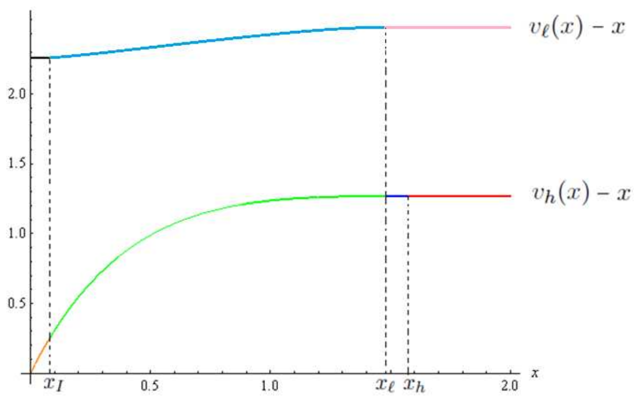

Figure 5.5: Value function with parameters in (5.4.5)

5.4.3 Different Cost but Same Interest Rate

In the example above, the possibility to issue new equity in the good state allows to reverse the ranking of the thresholds. So, even if the opportunity cost of cash is lower ($r_l < r_h$) the firm will issue dividends more often in the good state. In order to understand the impact of issuing costs, we now study this particular case to understand the affect of the cost alone. Indeed, let

$$r(i) = r > 0, \qquad i = 0, 1, \qquad \gamma(0) \leq \gamma(1). \tag{5.4.6}$$

It is clear that when both $\gamma(0)$ and $\gamma(1)$ are very large, then there will not be any issuance and the problem is the same as the one studied in [3]. In fact, we have an easy quantification of this statement. Let $V(x, r)$ be the Jeanblanc-Picqué & Shiryaev value function defined in (5.2.3). Let $x^*(r)$ be the dividend payment threshold for this problem and set

$$\gamma^*(r) := V(x^*(r), r) - x^*(r).$$

Lemma 5.4.2. Assume (5.4.6). Then, new equity issues are never optimal and $v(x,i) = V(x,r)$, if and only if

$$\gamma(i) \geq \gamma^*(r), \qquad i = 0, 1.$$

Proof. Since V is concave, we directly verify that for every $x, \xi \geq 0$ and $i = 0, 1$,

$$\begin{aligned} V(x+\xi, r) - V(x,r) &\leq V(\xi, r) - V(0,r) = V(\xi, r) \\ &< \xi + \gamma^* \leq \xi + \gamma(i). \end{aligned}$$

Using this it is straightforward to show that the value function $V(x,r)$ solves the dynamic programming equation (5.4.4). Hence by uniqueness $v = V$. In particular there are never new equity issues.

To prove the converse, assume that there are never new equity issues. Then, $v = V$ where V solves the dynamic programming equation (5.4.4). In particular,

$$V(x,r) \geq V(x+\xi, r) - \xi - \gamma(i),$$

for all $x, \xi \geq 0$ and $i = 0, 1$. We take $\xi = x^*(r)$ and $x = 0$ to conclude. $\square$

Based on the above result, we computed the value functions for the following parameter values

$$r(0) = r(1) = 0.05, \quad \lambda = 0.3, \quad \sigma = 0.25, \quad \mu = 0.18, \tag{5.4.7}$$

with two different issuance costs:

$$\gamma(0) = 0.1489 < \gamma^*(r) = 2.60748 << \gamma(1),$$

$$\gamma(0) = 0.7756 < \gamma^*(r) = 2.60748 << \gamma(1).$$

In both cases, we decreased $\gamma(0)$ from γ^*. In all examples, there is issuance as proved in Lemma (5.4.2). There are three critical thresholds:

$$0 \leq z_0 := \text{issuance threshold},$$

i.e., it is optimal to make an issuance whenever the cash reserves are less than or equal to z_0 and when we are in state $i = 0$. Numerically we observed that of relatively high values of $\gamma(0)$ (i.e, values less than but close to γ^*), $z_0 = 0$. However, $z_0 > 0$ for sufficiently small values of $\gamma(0)$. Hence, there is a balance between the probability of going to a bad state in which issuance is too costly and the probability of recovery.

The other common features of the numerical results is that the dividend payment threshold $x(i)$ is smaller in the "good" state of the economy, i.e., we always find:

$$x(0) < x(1).$$

In other words, dividend payment starts at lower cash reserves when the economy is in a good state.

Below are the tables of these results and two representative graphs. In the first graph $z_0 > 0$ and the black curve is the issuance part. In the second $z_0 = 0$. In both graphs red parts correspond to the dividend payment region.

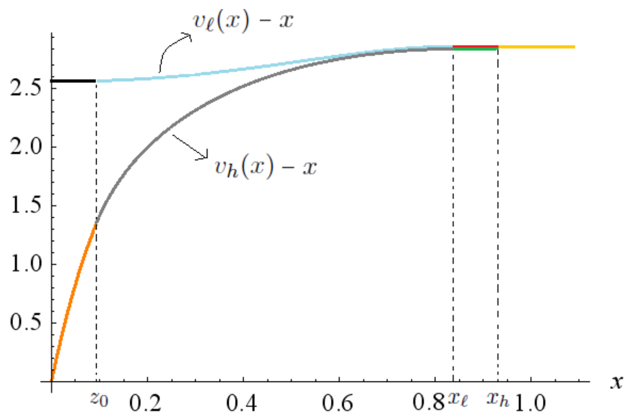

Figure 5.6: Value function with parameters in (5.4.7) and $\gamma(0) = 0.1489$

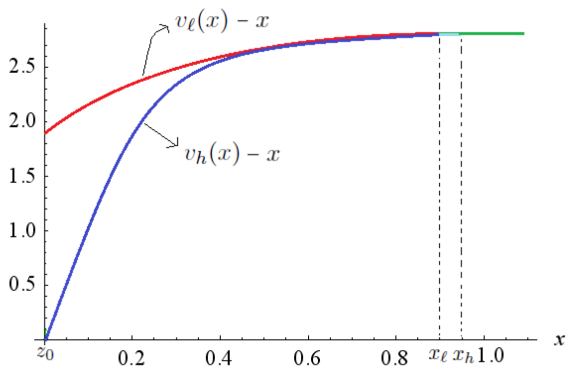

Figure 5.7: Value function with parameters in (5.4.7) and $\gamma(0) = 0.7756$

Table 5.1: Optimal values for the set of parameters $\sigma = 0.25$, $\mu = 0.18$, $r = 0.05$, $\lambda = 0.3$.

$\gamma(0)$	z_0	x_l	x_h
0.0002	0.4990	0.6726	0.9226
0.0033	0.3958	0.7229	0.9229
0.1236	0.1153	0.8327	0.9327
0.1490	0.0954	0.8390	0.9340
0.2691	0.0286	0.8582	0.9382
0.7756	0	0.9003	0.9503
1.0087	0	0.9159	0.9559
1.6265	0	0.9504	0.9704
2.0527	0	0.9702	0.9802

5.5 Conclusion

This paper has studied the specific impact of macroeconomic variables on the dividend policies of firms by considering the extreme case of a firm whose profitability is constant, but evolves in a stochastic macroeconomic environment, where interest rates and/or issuance costs are governed by an exogenous Markov chain.

Interestingly, we show that these two variables have opposed effects on the dividend policies of firms. Specifically, firms tend to distribute more dividends when interest rates are high and less dividends when issuing costs are high. We also find that stochastic issuing costs allow to get rid of the unfortunate prediction of previous models to which firms wait until the last moment (i.e. until they run out of cash) to issue new equity. Like Bolton, Chen & Wang [1], we obtain a market timing effect: when issuing costs are very high during recessions (so that shareholders refuse to recapitalize firms when they run out of cash) it becomes optimal to issue new equity in the good state even if the firm still has cash reserves, due to the fear that a recession might occur, leading to the forced closure of a profitable company.

Bibliography

[1] Bolton, P., Chen, H. and Wang, N. (2011). Market Timing, Investment and Risk Management. Working paper, Columbia University.

[2] Cadenillas, A. and Sotomayor, L.R. (2008). Classical Singular and Impulse Stochastic Control for the Optimal Dividend Policy when there is Regime Switching. Working paper, GSU and University of Alberta.

[3] Décamps, J-P, Mariotti, T., Rochet, J-C., and Villeneuve, S. (2011). Free Cash-Flow, Issuance Costs and Stock Price Volatility, Journal of Finance, 66(5), 1501 – 1544.

[4] Fleming, W.H., Soner, H.M. (1993). *Controlled Markov Processes and Viscosity Solutions*. Applications of Mathematics 25. Springer-Verlag, New York.

[5] Gertler, M., and Hubbard, R.G. (1993). Corporate Financial Policy, Taxation and Macroeconomic Risk. The RAND Journal of Economics, 24(2), 286 – 303.

[6] Hackbarth, D., Miao, J. and Morellec, E. (2006). Capital Structure, Credit Risk and Macroeconomic Conditions. Journal of Financial Economics, 82, 519 – 550.

[7] Karatzas, I., and S.E. Shreve (1991). *Brownian Motion and Stochastic Calculus, New York*, Springer.

[8] Jeanblanc-Picqué, M., and A.N. Shiryaev (1995). Optimization of the Flow of Dividends, Russian Mathematical Surveys, 50(2), 257 – 277.

[9] Jiang, Z., Pistorius, M. (2012). Optimal Dividend Distribution under Markov Regime Switching. Finance and Stochastics, 16(3), 449 – 476.

[10] Radner, R. and Shepp, L. (1996). Risk vs Profit Potential: A Model for Corporate Strategy. Journal of Economic Dynamics and Control, 20, 1373 − 1393.

[11] Rockefellar, R.T. 1970. *Convex Analysis*, Princeton University Press.

Appendix A

In this Appendix, we prove the concavity of the value function. Firstly, in view of Lemma (5.3.1) and Corollary (5.3.2), there are constants $c_1, c_2 > 0$ such that

$$v(x,i) \geq x + c_1 \quad \forall\, x \geq x^*/2,\ i \in \{0,1\} \tag{5.5.1}$$

$$v(x,i) \leq V(x, r_\ell) \leq x + c_2 \quad \forall\, x \geq 0,\ i \in \{0,1\}. \tag{5.5.2}$$

The following technical result is needed in the proof of concavity. Let x^* be as in the previous result. Also recall that $\theta^{x,L}$ is the exit time of $X^{x,L}$ from the interval $(0,\infty)$.

Lemma. There are $\hat{T} \geq 1$ and $\hat{\Lambda} < 1$ such that

$$\mathbb{E}[\Lambda_{\hat{T}\wedge\theta^{x,L}}] \leq \hat{\Lambda},$$

for all $x \geq x^*/2$, $L \in \mathcal{A}(x)$ satisfying

$$J(x,i;L) \geq x + \frac{c_1}{2},$$

where c_1 is as in (5.5.1).

Proof. Fix x and L as in the statement and set $X = X^{x,L}$. For $T > 0$ to be determined, set $\theta = \theta^{x,L}$ and $\tau := \theta \wedge T$. By dynamic programming,

$$J(x,i,L) \leq \mathbb{E}\left[\int_0^\tau \Lambda_t dL_t + \Lambda_\tau v(X_\tau, i_\tau)\right].$$

Set $\tilde{X}_t = x + \mu t + \sigma W_t$, so that $X_t = \tilde{X}_t - L_t$. Since $\Lambda_t \leq 1$, (5.5.2) implies

$$\begin{aligned} J(x,i,L) &\leq \mathbb{E}\left[\int_0^\tau dL_t + \chi_{\{\theta\geq T\}}(\tilde{X}_T - L_T + c_2)e^{-r_\ell T}\right] \\ &= \mathbb{E}\left[L_\tau\left(1 - \chi_{\{\theta\geq T\}}e^{-r_\ell T}\right) + \chi_{\{\theta\geq T\}}(\tilde{X}_T + c_2)e^{-r_\ell T}\right]. \end{aligned}$$

On $\{\theta < T\}$, $L_\theta = \tilde{X}_\theta$ and on $\{\theta \geq T\}$, we have $\tau = T$ and $L_T = \tilde{X}_T - X_T$. Then, since

$J(x,i;L) \geq x + c_1/2$,

$$\begin{aligned}
x + \frac{1}{2}c_1 &\leq J(x,i;L) \\
&\leq \mathbb{E}\left[\tilde{X}_\theta\ \chi_{\{\theta<T\}} + \left(\tilde{X}_T - X_T + e^{-r_\ell T}(X_T + c_2)\right)\chi_{\{\theta\geq T\}}\right] \\
&= \mathbb{E}\left[\tilde{X}_\tau + \left(-X_T + e^{-r_\ell T}(X_T + c_2)\right)\chi_{\{\theta\geq T\}}\right] \\
&= \mathbb{E}\left[\tilde{X}_\tau + \left(e^{-r_\ell T}c_2 - X_T(1 - e^{-r_\ell T})\right)\chi_{\{\theta\geq T\}}\right] \\
&\leq \mathbb{E}\left[\tilde{X}_\tau + e^{-r_\ell T}c_2\ \chi_{\{\theta\geq T\}}\right] \leq (x + \mu\mathbb{E}[\tau]) + e^{-r_\ell T}c_2.
\end{aligned}$$

We now set $T = \hat{T}$ where $\hat{T}$ is so that $e^{-r_\ell \hat{T}}c_2 = \frac{c_1}{4}$. Then,

$$x + \frac{c_1}{2} \leq x + \mu\mathbb{E}(\tau) + \frac{c_1}{4}.$$

Hence,

$$\mathbb{E}[\theta^{x,L} \wedge \hat{T}] = \mathbb{E}[\tau] \geq \frac{c_1}{4\mu}.$$

Set $f(t) = e^{-r_\ell t}$ so that $\Lambda_t \leq f(t)$. Since f is convex and $f(0) = 1$,

$$\begin{aligned}
\mathbb{E}[\Lambda_\tau] &\leq \mathbb{E}[f(\tau)] \leq \mathbb{E}\left[\frac{\tau}{\hat{T}}f(\hat{T}) + (1 - \frac{\tau}{\hat{T}})f(0)\right] \\
&= \frac{f(\hat{T})}{\hat{T}}\mathbb{E}[\tau] + \left(1 - \frac{1}{\hat{T}}\mathbb{E}[\tau]\right) \\
&= 1 - \frac{1}{\hat{T}}(1 - f(\hat{T}))\mathbb{E}[\tau] \\
&\leq 1 - \frac{1}{\hat{T}}(1 - f(\hat{T}))\frac{c_1}{4\mu} =: \hat{\Lambda}.
\end{aligned}$$

□

We are now ready to prove the concavity.

Proof of Theorem (5.3.3)**.** For $x, y \geq 0$, $i \in \{0,1\}$, set

$$I(x,y,i) := v(x,i) + v(y,i) - 2v\left(\frac{x+y}{2}, i\right).$$

In view of Corollary (5.3.1), $I(x,y,i) \le 0$, for all $x, y \in [0, x^*]$. Set

$$\hat{\alpha} := \sup \left\{ I(x,y,i) : \ \frac{x^*}{2} \le x \le y,\ i = 0,1 \right\}.$$

By Lemma (5.2.3), $\hat{\alpha} < \infty$. Hence, for every $\varepsilon > 0$ there are $x_\varepsilon, y_\varepsilon,\ i_\varepsilon \in \{0,1\}$ such that

$$\hat{\alpha} \le I(x_\varepsilon, y_\varepsilon, i_\varepsilon) + \varepsilon, \quad \text{and} \quad \frac{x^*}{2} \le x_\varepsilon \le y_\varepsilon.$$

In view of Lemma (5.3.1), to prove the concavity of v, it suffices to show that $\hat{\alpha} \le 0$.

Let $L^x \in \mathcal{A}(x_\varepsilon)$, $L^y \in \mathcal{A}(y_\varepsilon)$ be arbitrary dividend strategies satisfying

$$J(x_\varepsilon, i; L^x) \ge x_\varepsilon + \frac{c_1}{2}, \qquad J(y_\varepsilon, i; L^y) \ge y_\varepsilon + \frac{c_1}{2}. \tag{5.5.3}$$

In view of (5.5.1), such processes exist, and

$$v(x_\varepsilon, i) = \sup\{J(x_\varepsilon, i; L^x) \mid L^x \in \mathcal{A}(x_\varepsilon) \text{ and } L^x \text{ satisfies (5.5.3) }\}.$$

The same also holds at y_ε. Set

$$\bar{L} := \frac{L^x + L^y}{2}, \qquad \bar{x} := \frac{x_\varepsilon + y_\varepsilon}{2}.$$

Finally, let $\hat{T}$ be as in the Lemma above. Set $\theta^x := \theta^{x_\epsilon, L^x}$. Without loss of generality assume that

$$X_t^\varepsilon := x_\varepsilon + \mu t + \sigma W_t - L_t^x \le Y_t^\varepsilon := y_\varepsilon + \mu t + \sigma W_t - L_t^y, \ \ \forall\ t \le \theta^x.$$

Otherwise, one may simply redefine L^x and L^y so that $X_t^\varepsilon = Y_t^\varepsilon$ after the first time they are equal.

Set $\tau := \theta^x \wedge \hat{T}$. By the dynamic programming principle (5.2.7),

$$\begin{aligned}
J(x_\varepsilon, i; L^x) + J(y_\varepsilon, i; L^y) &\leq 2\mathbb{E}\left[\int_0^\tau \Lambda_t d\bar{L}_t\right] + \mathbb{E}\left[\Lambda_\tau(v(X^\varepsilon_\tau, i_\tau) + v(Y^\varepsilon_\tau, i_\tau))\right] \\
&= 2\mathbb{E}\left[\int_0^\tau \Lambda_t d\bar{L}_t + \Lambda_\tau v(X^{\bar{x},\bar{L}}_\tau, i_\tau)\right] \\
&\quad + \mathbb{E}\left(\Lambda_\tau\left[v(X^\varepsilon_\tau, i_\tau) + v(Y^\varepsilon_\tau, i_\tau) - 2v(X^{\bar{x},\bar{L}}_\tau, i_\tau)\right]\right) \\
&\leq 2v(\bar{x}, i) + \mathbb{E}[\Lambda_\tau]\hat{\alpha}.
\end{aligned}$$

By the Lemma above, $\mathbb{E}[\Lambda_\tau] \leq \hat{\Lambda} < 1$. Also,

$$v(x_\varepsilon, i_\varepsilon) + v(y_\varepsilon, i_\varepsilon) = \sup\{J(x_\varepsilon, i_\varepsilon; L^x) + J(y_\varepsilon, i_\varepsilon; L^y) \mid (L^x, L^y) \text{ satisfying (5.5.3)}\}.$$

Hence,

$$v(x_\varepsilon, i_\varepsilon) + v(y_\varepsilon, i_\varepsilon) \leq 2v(\bar{x}, i_\varepsilon) + \hat{\Lambda}\hat{\alpha}.$$

By the choice of $(x_\varepsilon, y_\varepsilon)$,

$$\hat{\alpha} \leq v(x_\varepsilon, i_\varepsilon) + v(y_\varepsilon, i_\varepsilon) - 2v(\bar{x}, i_\varepsilon) + \varepsilon \leq \hat{\Lambda}\hat{\alpha} + \varepsilon.$$

Hence $\hat{\alpha} \leq \varepsilon/(1 - \hat{\Lambda})$, for all $\varepsilon > 0$. Therefore, $\hat{\alpha} \leq 0$ and consequently v is concave.

$\square$

Appendix B

Proof of Proposition (5.3.5)**.** Towards a contradiction, suppose that $x_\ell < x_h$. Set

$$u(x) := v'_\ell(x), \quad w(x) := v'_h(x), \qquad \lambda_\ell := \lambda(0), \quad \lambda_h := \lambda(1).$$

Differentiating the original system once and using the above definitions yield the following coupled ordinary differential equations for u and w, on the interval $(0, x_\ell)$,

$$r_h w(x) = \mu w'(x) + (1/2)\sigma^2 w''(x) - \lambda_h[w(x) - u(x)], \tag{5.5.4}$$

$$r_\ell u(x) = \mu u'(x) + (1/2)\sigma^2 u''(x) + \lambda_\ell[w(x) - u(x)]. \tag{5.5.5}$$

Since $v_\ell(0) = v_h(0) = 0$ and $v_\ell(x) \geq v_h(x)$ for all $x \in [0, \infty)$, we conclude that $u(0) \geq w(0)$.

Our goal is to show that $u(x) \geq w(x)$ for all $x \in [0, x_\ell]$. Indeed, by our hypothesis $x_\ell < x_h$, $w(x_\ell) > w(x_h) = 1$. So if we can prove that $u \geq w$ on $[0, x_\ell]$, then

$$1 = u(x_\ell) \geq w(x_\ell) > 1$$

will provide the desired contradiction. Hence it suffices to prove that $u \geq w$ on $[0, x_\ell]$.

Set $\Phi(x) = (u - w)(x)$ and choose $y \in [0, x_\ell]$ so that

$$(u - w)(y) = \min_{x \in [0, x_\ell]} (u - w)(x) =: \alpha. \tag{5.5.6}$$

Our goal is to show that $\alpha \geq 0$. We analyze three cases separately.

Case 1: $y = 0$. In this case, $\alpha = u(0) - w(0) = 0$.

Case 2: $y \in (0, A)$. Since y is a local minimum of Φ,

$$\Phi'(y) = u'(y) - w'(y) = 0, \qquad \Phi''(y) = u''(y) - w''(y) \geq 0.$$

We use these first in (5.5.4) and then in (5.5.5) at the point y. The result is the following,

$$\begin{aligned} r_\ell u(y) &= \mu u'(y) + \frac{1}{2}\sigma^2 u''(y) - \lambda_\ell \alpha \geq \mu w'(y) + \frac{1}{2}\sigma^2 w''(y) - \lambda_\ell \alpha \\ &= r_h w(y) - [\lambda_h + \lambda_\ell]\alpha \geq r_\ell w(y) - [\lambda_h + \lambda_\ell]\alpha. \end{aligned}$$

In the a last step we used the fact that $w \geq 0$. Since $\alpha = u(y) - w(y)$, the above implies that $\alpha \geq 0$.

Case 3: $y = A$. By the smooth fit, we know that $v''(x_\ell) = u'(x_\ell) = 0$. We directly conclude that

$$\Phi'(x_\ell) = u'(x_\ell) - w'(x_\ell) = v_\ell''(x_\ell) - v_h''(x_\ell) = -v_h''(x_\ell) \geq 0.$$

Since $y = x_\ell$ is the minimum of Φ on the interval $[0, x_\ell]$, $\Phi'(x_\ell) \leq 0$. Hence, $\Phi''(x_\ell) = -v_h'(x_\ell) = 0$.

Recall that we have assumed that $x_h > x_\ell$. Set $f(x) := v_h''(x)$ and differentiate the dynamic programming equation (5.2.6) for v_h twice. The result is,

$$r_h f(x) = \mu f'(x) + \frac{1}{2}\sigma^2 f''(x) - \lambda_h f(x), \quad x \in (x_\ell, x_h),$$

together with boundary conditions $f(x_\ell) = f(x_h) = 0$. However, the zero function is the unique solution of this equation. Hence, $f(x) = v_h''(x) = 0$ for $x \in [x_\ell, x_h]$. So, v_h' is constant on $[x_\ell, x_h]$ as well. Since $v_h'(x_\ell) > 1$, we conclude that $x_h = \infty$. But this implies that $v_h(x) > v_\ell(x)$ for all sufficiently large x.

Hence, $x_\ell \geq x_h$.

□

Chapter 6

Erdinç Akyıldırım - Curriculum Vitae

Date / Place of Birth : 05.01.1982 / Turkey

Educational Background

2009 - 2013	**University of Zurich, Switzerland**
	Swiss Finance Institute PhD Program in Finance
	Advisors: Prof. Dr. H. Mete Soner and Prof. Dr. E. Walter Farkas
2005 - 2008	**Boğaziçi University, Turkey**
	Master of Science in Financial Engineering
2000 - 2005	**Boğaziçi University, Turkey**
	Bachelor of Science in Mathematics

Work Experience

2010-2013	**Mathematics Department, ETH Zurich, Switzerland**
	Research & Teaching Assistant
2008-2009	**Faculty of Engineering and Natural Sciences, Sabancı University, Turkey**
	Research & Teaching Assistant
2005-2006 & 2007-2008	**Mathematics Department, Boğaziçi University, Turkey**
	Research & Teaching Assistant
2006-2007	**Structured Products and Derivatives Desk,**
	Industrial Development Bank of Turkey
	Quantitative Analyst

Awards and Scholarships

- Grant provided by Swiss Finance Institute during the first year of doctoral studies
- Selected by Industrial Development Bank of Turkey as a member of a 13 people management trainee team among 3000 candidates.
- 2^{nd} rank in Mathematics Department graduate class.
- 1999-2001 The Ministry of Education Fellowship
- Among the the first 1500 entrants in the Nationwide University Entrance Exam among 1.5 million participants .
- 3^{rd} rank in high school graduate class.